

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 678.508 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_1_400

Perfect score: 2143

Sequence: 1 MDVKRPYSLTERRDAER.....EITEDTASSWPFVDVSLYP 400

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2_1/USPTO.spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CGN 1_1368 @runat_06082004_112215_29265 -NCFU=6 -ICFU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Geneseq 29Jan04:*
- 2: Geneseq 1980s:*
- 3: Geneseq 1990s:*
- 4: Geneseq 2000s:*
- 5: Geneseq 2001as:*
- 6: Geneseq 2001bs:*
- 7: Geneseq 2002as:*
- 8: Geneseq 2003bs:*
- 9: Geneseq 2003cs:*
- 10: Geneseq 2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|------------|--------------------|
| 1 | 2143 | 100.0 | 8354 | 6 ABS52100 | Abs52100 Human TEN |
| 2 | 2138 | 99.8 | 3111 | 2 AAV19251 | AAV19251 Human gam |
| 3 | 2138 | 99.8 | 3111 | 2 AAX87705 | AAX87705 Gamma-her |
| 4 | 2138 | 99.8 | 3111 | 3 AAD00791 | AAD00791 Human Her |
| 5 | 2138 | 99.8 | 3111 | 4 AAS18526 | AAS18526 DNA encod |
| 6 | 2098 | 97.9 | 8438 | 6 ABR85378 | ABR85378 Human NOV |
| 7 | 2082 | 97.2 | 8645 | 6 ABS78652 | ABS78652 Human cDN |
| 8 | 1285 | 60.0 | 2387 | 2 AAV19252 | AAV19252 Human gam |

| | | | | | |
|----|-------|------|-------|------------|--------------------|
| 9 | 987 | 46.1 | 1534 | 9 ADE07179 | Ade07179 Novel cod |
| 10 | 987 | 46.1 | 8645 | 6 ABQ82344 | Abq82344 Human NOV |
| 11 | 987 | 46.1 | 8675 | 6 ABQ82343 | Abq82343 Human NOV |
| 12 | 948 | 44.2 | 9695 | 7 ACC72052 | ACC72052 BCU0205B |
| 13 | 839 | 39.2 | 9729 | 5 AAS14089 | Aas14089 Human FCT |
| 14 | 839 | 39.2 | 9729 | 5 ADB32028 | Adb32028 Human FCT |
| 15 | 839 | 39.2 | 9826 | 5 AAS14085 | Aas14085 Human FCT |
| 16 | 839 | 39.2 | 9826 | 9 ADB32023 | Adb32023 Human FCT |
| 17 | 742.5 | 34.6 | 13202 | 4 AAK51828 | Aak51828 Human pol |
| 18 | 734.5 | 34.3 | 12879 | 6 ABK92230 | Abk92230 Prostate |
| 19 | 653.5 | 30.5 | 8473 | 6 ABQ82345 | Abq82345 Human NOV |
| 20 | 653.5 | 30.5 | 8487 | 6 ABQ82346 | Abq82346 Human NOV |
| 21 | 539 | 25.2 | 3052 | 6 ABQ75971 | Abq75971 Human pai |
| 22 | 534 | 24.9 | 1755 | 4 AAF23418 | Aaf23418 Human SEC |
| 23 | 533.5 | 24.9 | 978 | 9 ADA53142 | Ada53142 Human cod |
| 24 | 508 | 23.7 | 1727 | 7 ACC72051 | Acc72051 BCU0205A |
| 25 | 508 | 23.7 | 9058 | 8 ACD40264 | Acd40264 Human bre |
| 26 | 434 | 20.3 | 1429 | 5 AAS14084 | Aas14084 Human FCT |
| 27 | 434 | 20.3 | 1430 | 5 ADB32021 | Adb32021 Human FCT |
| 28 | 434 | 20.3 | 1431 | 4 AAF27861 | Aaf27861 Human NOV |
| 29 | 434 | 20.3 | 1431 | 4 AAS01213 | Aas01213 DNA encod |
| 30 | 434 | 20.3 | 1743 | 5 AAS67421 | Aas67421 DNA encod |
| 31 | 404 | 18.9 | 2676 | 5 AAS67421 | Aas67421 DNA encod |
| 32 | 344.5 | 16.1 | 452 | 6 ABS72992 | Abs72992 Human gen |
| 33 | 279 | 13.0 | 328 | 2 AAV87144 | Aav87144 EST clone |
| 34 | 279 | 13.0 | 1692 | 7 ACD05836 | Acd05836 Novel hum |
| 35 | 277 | 12.9 | 643 | 5 AAS67419 | Aas67419 DNA encod |
| 36 | 167.5 | 7.8 | 784 | 5 AAS69256 | Aas69256 DNA encod |
| 37 | 142.5 | 6.6 | 24533 | 4 AAS27689 | Aas27689 DNA encod |
| 38 | 142.5 | 6.6 | 24533 | 7 ABT17017 | Abt17017 Human sec |
| 39 | 142.5 | 6.6 | 24533 | 7 ABZ68157 | Abz68157 Human sec |
| 40 | 142.5 | 6.6 | 24533 | 7 ABZ74637 | Abz74637 Secreted |
| 41 | 142.5 | 6.6 | 24533 | 9 ADB94492 | Adb94492 Novel hum |
| 42 | 141 | 6.6 | 9222 | 7 AAL61171 | Aal61171 Actinosyn |
| 43 | 141 | 6.6 | 9222 | 7 AAL61224 | Aal61224 Actinosyn |
| 44 | 140 | 6.5 | 6034 | 9 ADD29677 | Add29677 Human tum |
| 45 | 140 | 6.5 | 6131 | 6 ABZ11281 | Abz11281 Human pol |

ALIGNMENTS

RESULT 1
ABS52100
ID ABS52100 standard; DNA; 8354 BP.
XX
AC ABS52100;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human TEN-M4-like gene.
XX

Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis; cell signal processing; metabolic pathway modulation; metabolic disorder; obesity; diabetes; infectious disease; neurodegenerative disorder; acne; Alzheimer's disease; Parkinson's disease; immune disorder; cancer; haematopoietic disorder; cirrhosis; pancreatitis; learning defect; memory defect; infertility; congenital heart defect; hair growth; pigmentation disorder; endocrine disorder; respiratory disease; health; gastro-intestinal disease; reproductive; neurological disease; bone marrow transplantation; endocrine disease; allergy; inflammation; nephrological disorder; urinary system disorder; age-related disorder; neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4; adipocyte complement-related C1q tumour necrosis factor; out at first; beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter; type Ia membrane sushi-containing domain; butyrophilin; single nucleotide polymorphism.

Homo sapiens.

Key Location/Qualifiers
FT variation replace(117,G) /*tag= a

variation /standard_name= "Single nucleotide polymorphism (SNP)"
 replace(225,C)
 /*tag= b
 variation /standard_name= "Single nucleotide polymorphism (SNP)"
 replace(260,A)
 /*tag= c
 /standard_name= "Single nucleotide polymorphism (SNP)"

WO200257453-A2.

25-JUL-2002.

19-DEC-2001; 2001WO-US050331.

19-DEC-2000; 2000US-0256704P.

20-DEC-2000; 2000US-0257314P.

02-MAY-2001; 2001US-0288153P.

29-MAY-2001; 2001US-0294075P.

24-JUL-2001; 2001US-0307506P.

10-AUG-2001; 2001US-0311590P.

10-AUG-2001; 2001US-0311613P.

29-AUG-2001; 2001US-0315617P.

14-SEP-2001; 2001US-0322358P.

(CURA-) CURAGEN CORP.

XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;

PI Stone DJ, Anderson D, Shinkets RA, Burgess CE, Zerhusen BD, Liu X;

PI Spytek KA, Casman SU, Boldog FL, Smithson G, Li L, Ji W;

DR WPI; 2002-590744/63.

DR P-PSDB; ABG70388.

XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy, atherosclerosis, metabolic disorders, diabetes, obesity, infectious disease, anorexia, neurodegenerative disorders, Alzheimer's disease or cancer.

XX Claim 8; Page 50-52; 318pp; English.

XX The present invention relates to new NOVX polypeptides. The invention is useful for treating or preventing a NOVX-associated disorder such as cardiomyopathy or atherosclerosis, where the disorder is related to cell signal processing and metabolic pathway modulation in a subject, preferably human. The invention is also useful for treating metabolic disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disease), immune disorders, haematopoietic disorders and various cancers. The molecules of the invention are also useful for treating or preventing cirrhosis, pancreatitis, learning and memory defects, infertility, congenital heart defects, acne, hair growth, pigmentation disorders, endocrine disorders, respiratory disease, gastro-intestinal diseases, reproductive diseases, neurological diseases, bone marrow transplantation, endocrine diseases, allergy and inflammation, nephrological disorders, urinary system disorders, neuropsychiatric disorders and age-related disorders. The present nucleic acid sequence represents a NOVX gene. This sequence encodes a NOVX protein of the invention

SQ Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 2,42e-127 | Length: | 8354 |
| Score: | 2143.00 | Matches: | 400 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| Ds: | 6 | Gaps: | 0 |

US-10-029-020-14_COPY_1_400 (1-400) x ABS52100 (1-8354)

Qy 1 MetAspValIysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20

Db 35 ATGGAGCTGAGGAGGAGGAGCCCTTACCGCTCGTACCCGCGCCGCGAGCCGAGCCG 94

| | | | |
|----|------|---|------|
| Qy | 21 | ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer | 40 |
| Db | 95 | CGCTACACAGCTCGTCCGCGGACAGCGAGGAGGCAAGCCCGCGAGAAATCGTACAGC | 154 |
| Qy | 41 | SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal | 60 |
| Db | 155 | TCCAGCGAGAGCCCTGAGGCGCTACGACGAGGAGCGCGCCTAGCCTATGCGAGCGCGTC | 214 |
| Qy | 61 | LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu | 80 |
| Db | 215 | AAGGACATTGTGCCGACGAGCGCGAGAAATCTCCGCGACAGGTGCCAACTTCACCTG | 274 |
| Qy | 81 | ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle | 100 |
| Db | 275 | CGGAGCTGGGGCTGAGAGAGTAGTACGCCCTTCACGGGACCTGTACCGGACAGACATT | 334 |
| Qy | 101 | GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp | 120 |
| Db | 335 | GGCCTGCCCAATGCGGCTACTCCATGGGGGTGGCTCTGTATGCCGACATGAGGCGTAC | 394 |
| Qy | 121 | ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg | 140 |
| Db | 395 | ACGGTGTCTGCCCTGAGCACCCCTGTGTGGGCGGAGCACACGGTCAGGGCCG | 454 |
| Qy | 141 | SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu | 160 |
| Db | 455 | AGCTCTCGCTCTCCAGCGCGGCCAATTCACACTCACCGACACCGGACGATGAA | 514 |
| Qy | 161 | AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro | 180 |
| Db | 515 | AAACTGAGACTGATCATCCGGCGGCGCTGCAGAACCCACGCGCGGTCCGCGCGCGG | 574 |
| Qy | 181 | ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu | 200 |
| Db | 575 | CGCGCGCTCTCCAGCGCCACACCCCAACACGACACCGCGGCTCCATTAATCCCTG | 634 |
| Qy | 201 | AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu | 220 |
| Db | 635 | AAACCGGGGCAACTTCACGCGGAGGAGCAACCCACGCGCGCCCGCCAGCACCTCGCTC | 694 |
| Qy | 221 | SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnAsnTrpLeu | 240 |
| Db | 695 | TCCGAGAGCGCCCTCGCGCGCGCGCCAGAGAGCTGCCCGACGCCAGAGAACTGGCTG | 754 |
| Qy | 241 | LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr | 260 |
| Db | 755 | CTCAACAGCAACATCCCTCTGGAGACAGAAACCTAGGCAGACGACCATTCCTAGGACA | 814 |
| Qy | 261 | LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr | 280 |
| Db | 815 | TTGCGAGGCAACCTCATTGAGATGGACATTCCTCGCGCGCTCCCGCCATCATGGGGCTTAC | 874 |
| Qy | 281 | SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer | 300 |
| Db | 875 | AGTGACGGGCACTTCCTCTCAAGCCTGGAGGACCTCCCGCTCTCTTCGCACCATCA | 934 |
| Qy | 301 | ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg | 320 |
| Db | 935 | CCAGGGTACCCACTGACGTCACGACAGTACTCTCTCCGCGCGCCCGCCCTGCCCCG | 994 |
| Qy | 321 | SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys | 340 |
| Db | 995 | AGCACCTTCGCGCGCGCGCCCTTTAACTCTCAAGAGCCCTCCCAAGTACTGTAACTGGAA | 1054 |
| Qy | 341 | CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe | 360 |
| Db | 1055 | TGGCAGCGCCCTGAGCGCCATCGTCACTCAGCACCTCTGCTCATCTCTGCTGCGTACTTT | 1114 |
| Qy | 361 | ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr | 380 |
| Db | 1115 | GTGGCCATGCACCTGTTTGGCTTAAACTGGCACCTGCAGCCGATGGAGGGGAGATGTAT | 1174 |


```

Db      1414 GTGGCCATGACCTCTTTGGCCCTAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1473
QY      381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db      1474 GAGATCAGGAGGACACAGCCAGCAGTTGGCCCTGTGCCAACCGACGCTCTCCCTATACCCC 1533

RESULT 3
AAx87705
ID   AAX87705 standard; cDNA; 3111 BP.
XX
AC   AAX87705;
XX
DT   26-OCT-1999 (first entry)
XX
DE   Gamma-hereregulin cDNA.
XX
KW   Gamma-hereregulin; HRG-alpha; human; ligand; HER2; HER3; HER4; receptor;
KW   lung surfactant; respiratory distress syndrome; emphysema;
KW   epithelial growth factor; therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key Location/Qualifiers
FT   CDS 334..2640
FT   FT /*tag= a
XX
PN   WO9939729-A2.
XX
PD   12-AUG-1999.
XX
PF   03-FEB-1999; 99MO-US002390.
XX
PR   04-FEB-1998; 98US-00020598.
XX
PA   (GETH ) GENENTECH INC.
PA   (IOWA ) UNIV IOWA RES FOUND.
XX
PI   Sliwkowski M, Kern JA;
XX
DR   WPI; 1999-494213/41.
XX
DR   P-PSDB; AAY06639.
XX
PT   Heregulin ligands can be used to induce epithelial cell growth, and to
PT   promote repair and healing of tissue damage or injury.
XX
PS   Disclosure; Page 111-116; 120pp; English.
XX
CC   This is the nucleotide sequence of cDNA coding for gamma-hereregulin (gamma
CC   -HRG, see AAY06639). The invention provides HRG ligands, including gamma-
CC   HRG, that have affinity for and stimulate HER2, HER3 and/or HER4
CC   receptors in autophosphorylation. A new method of treating respiratory
CC   distress syndrome in humans uses HER2, HER3 and/or HER4 receptor ligands
CC   as epithelial growth factors. A novel method of inducing epithelial cell
CC   growth and/or proliferation comprises contacting a normal epithelial cell
CC   which expresses HER2, HER3 and/or HER4 receptors with an isolated ligand
CC   which activates HER2, HER3, HER4 receptors or their combination. Also
CC   claimed are methods of increasing lung surfactant protein A, or of
CC   treating chronic obstructive pulmonary disease, respiratory distress or
CC   emphysema, by administering an effective amount of an isolated HER ligand
CC   to a patient
XX
SQ   Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,62e-127 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x AAX87705 (1-3111)

```

```

QY      1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db      334 ATGGACGTCGAGGAGAGGAGGAGCTTACCGCTCGCTGACCCGCGCGCGAGCCGCGAGCCG 393
QY      21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db      394 CGCTACACCGAGCTCGTCCGCGACAGCGAGGAGGGCAAGCCCGCAGAAATCGTACAGC 453
QY      41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrClySerArgVal 60
Db      454 TCCAGCGAGACCCCTGAGGCGCTTACACCGAGGAGCGCCCGCTAGCCTATATGGCAGCGCGTC 513
QY      61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db      514 AAGGACATTTGTCGCGAGGAGCGCGAGGAATTCGCGGCACAGGTGCCAACTTCACCTCG 573
QY      81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db      574 CGGAGAGCTGGGGCTGGAAGAGTAACCGCCCTCAACGGGACCTGTACCGGACAGACATT 633
QY      101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db      634 GGCCTCCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
QY      121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db      694 ACGTGTCTGCCCTGAGCACCCCGCTGCTGTGGGGCGCGAGCACACGGTCAGGGCGC 753
QY      141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db      754 AGCTCTGCTGTCCAGCGGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
QY      161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db      814 AACACTGAGACTGATCATCCGGCGGCTTGCAAGAACCCGCGGCTCCGAGCGCGCGCG 873
QY      181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db      874 CCGCGCTCTGCGACGCCACACCCCAACACGAGCACCGCGGCTCCATTAATCCCTG 933
QY      201 AsnArgGlyAsnPheThrProArgSerSerProSerProAlaProThrAspHisSerLeu 220
Db      934 AACCGGGGCAACTTCACGCGAGGAGCAACCCCGCGCGCGCGCGCGCGACCATCGCTC 993
QY      221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTyrLeu 240
Db      994 TCCGAGAGCCCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1053
QY      241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db      1054 CTCAACAGCAACATCCCCCTGGAGACCAAGAAACCTAGGCAAGCAGCGCATTCCTAGGGA 1113
QY      261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db      1114 TTTCAGGCAACCTCATTTGAGATGGACATTCCTCGCGCGCTCCCGCATGATCGGGCTTAC 1173
QY      281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db      1174 AGTGAGGCGACTTCTCTTCAAGCTGGAGGACCTCCCGCTCTTCTGCGCACCATCA 1233
QY      301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db      1234 CCAGGGGTACCCACTGACGTCCAGCACAGTGTAATCTCTCTCCCGCGCGCGCGCGCGCG 1293
QY      321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
Db      1294 AGCACTTCGCGCGCGCGCGCTTTAACTCAAGAGCCCTCCCAAGTACTGTAACTGGAAG 1353
QY      341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db      1354 TGGCAGCGCTCGAGCGCATCGTCATCTCAGCCACTCTGTCATCTCTGTCGTCATCTTT 1413

```

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGCCCATGACCTGGTGGCTAAACTGGCACTGCAGCCGATGGAGGGCGCAGATGAT 1473
QY 381 GlnIleThrGluAspThrAlaSerSerTyrProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCAGCAGTGGCTGTGCCACCGACGCTCCCTATATACCC 1533

RESULT 4
AAD00791
ID AAD00791 standard; cDNA; 3111 BP.
AC AAD00791;
XX 21-SEP-2000 (first entry)
DE Human Heregulin variant, gamma-HRG encoding cDNA.
XX Heregulin; variant; gamma-HRG; human; inner-ear-supporting cell;
KW activator; HER2 receptor; HER3 receptor; HER-2 oncogene; chromosome 8p;
KW hair cell related hearing disorder; ototoxic injury; tissue damage;
KW acoustic assault; degenerative hearing loss; balance impairment;
KW treatment; hair cell; surgical injury; physical injury;
KW inner ear disorder; ss.
XX Homo sapiens.

OS
XX Key Location/Qualifiers
FH 334..2640
CDS /*tag= a
FT /product= "Human Heregulin variant, gamma-HRG protein"
FT
FT
XX WO200027426-A1.

XX 18-MAY-2000.
XX 28-OCT-1999; 99WO-US025744.
XX 07-NOV-1998; 98US-0107522P.
XX (GETH) GENENTECH INC.
XX Gao W;
XX WPI; 2000-376313/32.
DR P-PSDB; AAY711176.

XX Method for inducing hair cell generation and inner-ear-supporting cell
PT growth regeneration and proliferation, useful for treating hearing
PT disorders.
XX Disclosure; Fig 7; 141pp; English.

XX The patent discloses a method for inducing hair cell generation, or inner
CC -ear-supporting cell growth, regeneration, and/or proliferation, by
CC heregulin (HRG), the ligand for HER2/HER3 dimeric receptors. Heregulin
CC proteins function as activators of HER-2 oncogene and result from
CC alternate splicing of a single gene mapped to chromosome 8p. The two
CC major types, alpha and beta HRG's are based on two variant EGF-like
CC (epidermal growth factor) domains, that differ in their C-terminal ends.
CC This method can be used to increase the number of inner-ear-supporting
CC cells and for treatment of hair cell related hearing disorders and
CC disease states associated with tissue damage, e.g. ototoxic injury,
CC acoustic assault, degenerative hearing loss, balance impairments, damage
CC associated with surgery or physical injury and inner ear disorders
CC related to hair cell dysfunction. The present sequence is the human
CC heregulin variant, gamma-HRG protein encoding cDNA, isolated from a
CC native HRG source, such as human MDA-MB-175 cells. It has at least one
CC biological property as the native sequence
XX
SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;
Alignment Scores:

Pred. No.: 1.62e-127 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 3 Gaps: 0
US-10-029-020-14_COPY_1_400 (1-400) x AAD00791 (1-3111)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGACGCTGAAGGAGAGAAAGCCCTTACCGCTCGCTGACCGCGCGCGAGCGCGC 393
QY 21 ArgTyrThrSerSerSerAlaAspSerGluGluLysAlaProGlnLysSerTyrSer 40
Db 394 CGTACACCAAGCTCGTCCGCGGACAGCGAGGAGGCAAGCCCGCAAGAAATCGTAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCCCTGAAGGCTTACGACCGAGACCGCCGCTAGCCTATGGCAGCGGTC 513
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 514 AAGGACATTCTCGCGAGGAGCGCGGAATTTCTGCGCACAGGTGCCAACTTCACCTG 573
QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGGCTGGAAAGTAACGCCCTCACGGGACCCCTGTACCGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyValacylSerAspAlaAspMetGluAlaAsp 120
Db 634 GGCTCCCCACCTGGGGCTACTTCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTGAGCACCGCGCTGTGTGGGGCGCGAGCACACCGTCAGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGTCTCTGCTGCCAGCGGGCCAAATTCAAATCTCACATCCCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGluAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGGGCGCTGCAGAACACCGCGGGCTCCGAGCGCGCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGGCTCTCGCACCGCCACACCCCCAACACAGCACCGCGGGCTCCATTACTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCACGCGGAGAGCAACCCAGCGCGGGCCCGCCCGACCGACCATC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCCCTCGCGCGCGCGCCAGAGGCTGCCACCGCCGAGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAGCAACATCCCTCGGAGAGCAGAAACCTAGGCAAGCAGGCTTCCTTAGGACA 1113
QY 261 LeuGlnAspAsnLeuLeuGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTCAGAGCAACTCTTATGATGGACATTCTCGGGCGCTCCCGCCCATGATGGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGACATTCCTCTTCAGACCTGGAGGACACCTCCCGGCTCTTCTGCACCATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProProLeuProArg 320
Db 1234 CCAGGGTACCACTGACGTCAGCACAGTGTACTCTCTCCCGCGCGCGACCCCTGCCCCCGC 1293

QY 201 AsnArgGlyAsnPhenThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
 DB 934 AACCGGGCAACTTACCGCGAGAGCAACCCAGCCCGCGCCACCGACCACTCGCTC 993
 QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
 DB 994 TCCGGAGAGCCCTGCGCGGGCGCCAGGAGCTGCCACGCCAGAGAGAACTGGCTG 1053
 QY 241 LeuAsnSerAsnLeuProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
 DB 1054 CTCACAGCAACATCCCTCCCTGGAGACAGAAACCTAGGCAAGCAGCCATCTTAGGGACA 1113
 QY 261 LeuGlnAsnLeuLeuGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
 DB 1114 TTGAGAGAACCTCATGTAGATGGACATTCTCGCGCGCTCCCGCCATGATGGGGCTTAC 1173
 QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
 DB 1174 AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCACTCA 1233
 QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
 DB 1234 CCAGGGTACCACTGACGTGCAGCACAGTGTACTCTCTCGCGCCGACCCCTGCCCGCG 1293
 QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
 DB 1294 AGCACCTTCGCCCGCGCGCTTTAACTCAAGAGCCCTCAAGAGTACTGTAACTGGAAG 1353
 QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
 DB 1354 TGGCAGCCCTGAGCGCATCGTCATCTCAGCCACTCTGGTCATCTGCTGGCATATTT 1413
 QY 361 ValAlaMetHisLeuPheGlyLeuAsnThrHisLeuGlnProMetGluGlyGlnMetTyr 380
 DB 1414 GTGGCCATGCACCTGTTGGCTAACTGGCACCTGCAGCGCATGTGAGGGGCGCATGTAT 1473
 QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
 DB 1474 GAGATCAGGAGACACAGCGAGCTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1533

RESULT 6
 ID ABN85378
 XX ABN85378 standard; DNA; 8438 BP.
 AC ABN85378;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human NOV1, TEN-M4 like protein, coding sequence.
 XX
 KW Human; NOV1; cytostatic; Cardiant; Antinflammatory; Immunosuppressive;
 KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
 KW Antisthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
 KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
 KW Gene Therapy; NOV; cancer; heart disease; inflammation;
 KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
 KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
 KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
 KW TEN-M4 like protein; chromosome 11; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 4..8395
 FT /*tag= a
 FT /trans except= (pos: 1138..1147,aa:Met)
 FT /product= "NOV1 protein"
 XX
 PN W020025704-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2002; 2002WO-US0000554.

XX 09-JAN-2001; 2001US-0260417P.
 PR 10-JAN-2001; 2001US-0260831P.
 PR 28-FEB-2001; 2001US-0272338P.
 PR 09-MAR-2001; 2001US-0274876P.
 PR 18-APR-2001; 2001US-0284704P.
 XX
 FA (CURA-) CURAGEN CORP.
 XX
 XX Padigaru M, Li L, Zerhusen BD, Casman SU, Shenoy S, Sptek KA,
 PI Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;
 PI Taylor S, Tchervet VT, Miller CE, Guo X, Boldog PL, Grosse WM;
 PI Alsbrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K,
 PI Macdougall J, Mayankar U, Millet I, Peyman J, Smithson G,
 PI Gunther E, Stone DJ;
 XX
 XX WPI; 2002-590674/63.
 DR P-PSDB; ABB98401.
 XX
 DR NOVX polypeptides and encoding polynucleotides, useful for preventing or
 PT treating NOVX-associated disorders e.g. cancer, inflammation, or
 PT Alzheimer's disease, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Claim 9; Page 8-9; 358pp; English.
 XX
 CC The present sequence is a coding sequence for a NOV protein. The NOV
 CC proteins and coding sequences are useful for treating or preventing NOV-
 CC associated disorders or in the manufacture of a medicament for treating
 CC the disorders, such as cancer, heart disease, inflammation, autoimmune
 CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
 CC IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
 CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
 CC and other wasting disorders associated with chronic diseases. NOV1 is a
 CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
 XX
 SQ Sequence 8438 BP; 2001 A; 2434 A; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.86e-124 Length: 8438
 Score: 2098.00 Matches: 398
 Percent Similarity: 98.76% Conservative: 0
 Best Local Similarity: 98.76% Mismatches: 2
 Query Match: 97.90% Indels: 3
 DB: Gaps: 1

US-10-029-020-14_COPY_1_400 (1-400) x ABN85378 (1-8438)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
 DB 4 ATGGACGTGAAGAGAGAGAAAGCCCTTACCGCTCGCTGACCCGCGCGCGAGCGGC 63
 QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
 DB 64 CGTACACCACTCGTCCGCGGACAGGAGGGCAAGCCCGCAGAGAAATCGTACAGC 123
 QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
 DB 124 TCCAGCGAGACCTCGAAGGCGCTACGACGAGCGCGCTAGCTATGGCAGCGCGTC 183
 QY 61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80
 DB 184 AAGACATTGTCCCGAGGAGCGCGAGAAATCTGCCGACAGGTGCCACTTCACCTG 243
 QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
 DB 244 CGGAGCTGGGGCTGGAGAGTAACGCCCCCTCCGAGGACCTGTACCGGACAGACATT 303
 QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
 DB 304 GGGCTCCCCACTGCGGTACTCATGGGGGCTGGCTCTCATGCCGACATGGAGGCTGAC 363
 QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140

```
Db 364 ACAGTGCTGCTCCCTGAGCACCCCGTGGTCTGTGGGGCGGAGCACACGGTCAGGGCGC 423
Qy 141 SerSerCysLeuSerSerArgAlaLenSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 424 AGCTCTGCTGCTCCAGCGGGCCAAATTCCAATCTCACACTCACCAGACACCGAGCATGAA 483
Qy 161 AsnThrGluThrAspHisProGlyLeuGluGlnAsnHisAlaArgLeuArgThrProPro 180
Db 484 AACACTGAGACTGATCATCCGGGGGGCTGGAGACCAGCGGGCTCCGGAGCGCGCGC 543
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 544 CGCGCGCTCTCGACGCCACACCCCAACACAGCACCGCGCGCTCCATTAACCTCCCTG 603
Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 604 AACCGGGGGAATTCACGCGGAGGAGCAACCCGCGCGCGCGCGCGCGCGCGCGCTC 663
Qy 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 664 TCCGGAGAGCCCTCGCGGGCGGCCAGAGGCTGCCACGCCAGGAGAACTGGCTG 723
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 724 CTCACAGCAACATCCCTCGGAGACCAAGAACTAGGCAAGCAGCATTCTTAGGACA 783
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 784 TTGCAGGAGCAACTATTGAGATGGACATTCGCGCGCGCTCCGCCATGATGGGGCTTAC 843
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 844 AGTGAGGGCACTTCTCTTCAGCTTGAGGACCTCCCGCTCTTCTGACCAATCA 903
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 904 CCAGGGTACCCACTGAGTCCAGCACAGGTACTCTCTCCCGCGGACCCCTGCCCGCG 963
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
Db 964 AGCACTTCTGCTGGCGGCTTTTAACTCAAGAGCCCTCCAGTACTGTAACCTGGAAG 1023
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1024 TCGCAGCCCTGAGCGCATGTCATCTCAGCCACTCTGGTCATCTCTGTCGCACTTT 1083
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet--- 379
Db 1084 GTGGCCATGCACTGTTGGCCCTAACTGGCACCTGACCCGATGGAGGGCGAGATGAGC 1143
Qy 380 ---TyrGluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLe 398
Db 1144 GATTATGATGATCAGGAGGACACAGCCAGTGTGGCTGTGCCAACCGACGTCTCCCT 1203
Qy 398 TyrPro 400
Db 1204 ATACCC 1210
RESULT 7
ABS78652
ID ABS78652 standard; cDNA; 8645 BP.
XX AC
XX ABS78652;
XX AC
DT 16-DEC-2002 (first entry)
XX
DE Human cDNA encoding CGDD10, INCYTE 7488573CB1.
XX
XX Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
XX cell proliferative disorder; arteriosclerosis; atherosclerosis;
XX cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
XX polycythaemia vera; primary thrombocytopaenia; developmental disorder;
```

```
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX WO200272830-A2.
XX 19-SEP-2002.
XX
XX 08-FEB-2002; 2002WO-US003715.
XX
XX 09-FEB-2001; 2001US-0268111P.
XX 23-FEB-2001; 2001US-0271175P.
XX 08-MAR-2001; 2001US-0274503P.
XX 09-MAR-2001; 2001US-0274552P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
XX Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
XX Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
XX
XX WPI; 2002-723356/78.
XX P-PSDB; ABG97359.
XX
XX New human proteins associated with cell growth, differentiation and
XX death, useful for diagnosing, treating or preventing autoimmune or
XX inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
XX atherosclerosis or hepatitis.
XX
XX Claim 5; Page 175-178; 181pp; English.
XX
XX The invention relates to an isolated polypeptide comprising CGDD1-12
XX (cell growth, differentiation and death), a naturally occurring amino
XX acid sequence at least 90% identical to CGDD, a biologically active
XX fragment or an immunogenic fragment. Also included are the
XX polynucleotides encoding CGDD1-12, a recombinant polynucleotide
XX comprising a promoter sequence operably linked to the CGDD
XX polynucleotides, a cell transformed with the recombinant polynucleotide, an anti-
XX a transgenic organism comprising the recombinant polynucleotide, or are
XX CGDD antibody, screening for compounds which bind to/modulate or are
XX ant/agonists of CGDD or alter the expression of CGDD polynucleotides and a
XX CGDD polynucleotide microarray. The polypeptides, polynucleotides,
XX agonists and antagonists are useful for diagnosing, treating or
XX preventing disorders associated with aberrant expression of CGDD,
XX particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
XX cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
XX vera, psoriasis, primary thrombocytopaenia or cancer), developmental
XX disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
XX neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
XX epilepsy), reproductive disorders (e.g. infertility or a disruption in
XX the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
XX (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
XX thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
XX glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
XX osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
XX uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
XX infections. They are also useful in the assessment of the effects of
XX exogenous compounds on the expression of nucleic acid and amino acid
XX sequences of proteins associated with CGDD. The present sequence encodes
XX a CGDD protein
XX
XX Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;
```

Alignment Scores:

Pred. No.: 2, 02e-123 Length: 8645

Score: 2082.00 Matches: 393

| | | | |
|------------------------|--------|---------------|---|
| Percent Similarity: | 98.25% | Conservative: | 0 |
| Best Local Similarity: | 98.25% | Mismatches: | 3 |
| Query Match: | 97.15% | Indels: | 4 |
| DB: | 6 | Gaps: | 2 |

| | | | |
|---|------|---|------|
| US-10-029-020-14_COPY_1_400 (1-400) x ABS78652 (1-8645) | | | |
| QY | 1 | MetAspValIysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg | 20 |
| DB | 117 | ATGACGTGAAGAGAGAACCTTACCGCTCGCTGACCCCGCGCCGCGAGCGC | 176 |
| QY | 21 | ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer | 40 |
| DB | 177 | GCCTACACCACTCGTCGCGGACAGCGAGGAGGCGCAAGCCCGCAGAAATCGTACAGC | 236 |
| QY | 41 | SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal | 60 |
| DB | 237 | TCCAGCGAGACCTGAAGGCTAGCAGCCAGGACCGCGCTAGCCTATGGCAGCGCGTC | 296 |
| QY | 61 | LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrIleu | 80 |
| DB | 297 | AAGACATTGTGCGCAGGAGGCGGAGAAATCTGCCGACAGGTGCCAACTTCACCCGTG | 356 |
| QY | 81 | ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle | 100 |
| DB | 357 | CGGAGCTGGGCTGGAGAGTAACGCCCTCAGCGGACCTGTACCGGACAGACATT | 416 |
| QY | 101 | GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp | 120 |
| DB | 417 | GGCTCCCGCCACTGCGGCTACTCCATCGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC | 476 |
| QY | 121 | ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg | 140 |
| DB | 477 | ACGGTGTGTTCCTTGAGACCCCGTGGCTGTGTGGGGCGGAGCACACGGTTCAGGGCGC | 536 |
| QY | 141 | SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu | 160 |
| DB | 537 | AGCTCTCGCTGTCCAGCGGGCGCAATTCCAATCTCACACTCCACCGACCGGACATGAA | 596 |
| QY | 161 | AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro | 180 |
| DB | 597 | AACACTGAGACT-----CGGGCGGGCTGCAGAACACCGCGCGCTCCGGACGCGCGC | 650 |
| QY | 181 | ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu | 200 |
| DB | 651 | CGCGCGCTCTCGACGCCCAACCCCCAACAGCACCGCGGCTCCATTAATCCCTG | 710 |
| QY | 201 | AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu | 220 |
| DB | 711 | AACCGGGGCAACTTCACGCGGAGGAGCAACCCCGAGCCCGGCCCCACCGACCTCGCTC | 770 |
| QY | 221 | SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu | 240 |
| DB | 771 | TCCGGAGAGCCCTGCGCGCGCGCGCCAGAGCGCTGCCCGCCAGGAGAACTGGCTG | 830 |
| QY | 241 | LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr | 260 |
| DB | 831 | CTCAACAGCAACATCCCTCTGGAGACAGAAACCTAGGCAAGACGACATTCCTAGGGACA | 890 |
| QY | 261 | LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr | 280 |
| DB | 891 | TTGCAGGACAACTCAATTGAGATGGACATTCTCGGGCGCTCCCGCCATGATGGGCTTAC | 950 |
| QY | 281 | SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer | 300 |
| DB | 951 | AGTGACGGGACATTCCTCTCTCAAGCTGGAGGACCTCCCGCTCTTCTGCACACATCA | 1010 |
| QY | 301 | ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg | 320 |
| DB | 1011 | CCAGGTACCACTGACCTCAGCACAGTGTACTCTCTCCGCGCCGACCCCTGCCCCCGC | 1070 |
| QY | 321 | SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys | 340 |

| | | | |
|----|------|--|------|
| Db | 1071 | AGCACCTTCGCCCGCGCGCTTAACTCAAGAGCCCTCAAGTACTGTAACTGGAAG | 1130 |
| QY | 341 | CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe | 360 |
| DB | 1131 | TGCGCAGCCCTGAGCGCATCGTCTCATCTCAGCCACTCTGGTCACTCTGCTGGCATACTT | 1190 |
| QY | 361 | ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr | 380 |
| DB | 1191 | GTGGGTAAAGCACCTCTTTC-----AACTGGCACCTGCAGCCGATGGAGGGCAGATGAT | 1244 |
| QY | 381 | GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro | 400 |
| DB | 1245 | GAGATCAGGAGGACACAGCCAGCAGTGGCTGTGCACACCGACGCTCTCCTTATACCC | 1304 |

| | | | |
|----------|---|------------|--|
| RESULT 8 | | | |
| AAV19252 | | | |
| ID | AAV19252 standard; cDNA; 2387 BP. | | |
| XX | AC AAV19252; | | |
| XX | 17-AUG-1998 (first entry) | | |
| DE | Human gamma-heretulin cDNA clone 20. | | |
| XX | Gamma-heretulin; gamma-HRG; human; autocrine growth factor; | | |
| KW | breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor; | | |
| KW | cell proliferation; cell differentiation; cell survival; | | |
| KW | neurological disorder; muscular disorder; ss. | | |
| OS | Homo sapiens. | | |
| XX | Key Location/Qualifiers | | |
| FT | CDS | 307..1890 | |
| FT | /*tag= a | | |
| FT | misc_feature | 1186..1263 | |
| FT | /*tag= b | | |
| FT | /note= "insert DNA" | | |
| XX | W09802541-A1. | | |
| XX | 22-JAN-1998. | | |
| XX | 08-JUL-1997; 97WO-US011841. | | |
| XX | 12-JUL-1996; 96US-0021640P. | | |
| XX | (GETH) GENENTECH INC. | | |
| XX | Schaefer GM, Sliwkowski M; | | |
| XX | WPI; 1998-110589/10. | | |
| XX | P-PSDB; AAW44818. | | |
| XX | DNA encoding gamma-heretulin - used to activate ErbB receptor and to | | |
| XX | enhance proliferation, differentiation or survival of a cell. | | |
| XX | Disclosure; Page 56-57; 81pp; English. | | |
| XX | Clone 20 encodes an N-terminally truncated isoform (see AAW44818) of | | |
| XX | human gamma-heretulin (gamma-HRG) (see also AAW44817), a novel member of | | |
| XX | the heretulin superfamily. It was isolated from a MDA-MB-175 human breast | | |
| XX | cancer cell library after screening with an EGF-like domain. Compared to | | |
| XX | gamma-HRG cDNA (see AAV19251), clone 20 contains a 78 base insert between | | |
| XX | codons 560 and 561, and lacks the 5' end. The truncated gamma-HRG was | | |
| XX | insoluble when expressed as a thioredoxin fusion protein in Escherichia | | |
| XX | coli cells. Gamma-HRG nucleic acid and polypeptide are claimed, together | | |
| XX | with their uses e.g. for enhancing the proliferation, differentiation or | | |
| XX | survival of glial or muscle cells | | |
| SQ | Sequence 2387 BP; 641 A; 640 C; 550 G; 556 T; 0 U; 0 Other; | | |

| | | | |
|-------------------|----------|---------|------|
| Alignment Scores: | 4.66e-73 | Length: | 2387 |
| Pred. No.: | | | |

```
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.96% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x AAV19252 (1-2387)

QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProProLeuSerHis 185
Db 1 CATCGGGCGGCTGTCAGAACACGCGGCTCCGAGCCGCCCGCGCTCTCGCAC 60

QY 186 AlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCACACACCCACACGACACGCGGCTCCATTAACCTCTGAAACCGGGGCACTTC 120

QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACGCCGAGGAGCAACCCACGCGGCCGCCACGACCTCTCTCGGAGAGCCCTC 180

QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuAsnSerHis 245
Db 181 GCCGGCGGGCCGAGAGCTGCCACGCCGAGGAACTGGCTGCTCAACAGCAACATC 240

QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
Db 241 CCCCTGGAGACCAAGAACTAGGACGAGCCATTCCTAGGACATTCGAGGACCACTC 300

QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTySerAspGlyHisPhe 285
Db 301 ATTGAGATGACATTCCTGGCGCTCCGCCATGATGGGGCTTACAGTCACGGGCACTTC 360

QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyProLeu 305
Db 361 CTCCTCAAGCCTGGAGGACCTCCCGCTCTCTGACCAACATCACAGGATACCCACTG 420

QY 306 ThrSerSerThrValTySerProProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 AGCTCCAGACAGTGACTCTCTCCGCCGCCGACCCCTGCCCGGACGACCTTCGCCCG 480

QY 326 ProAlaPheAsnLeuLysProSerLysTySerAsnTrpLysCysAlaAlaLeuSer 345
Db 481 CGGGCTTTAAGCTCAAGAGCCCTCCAAAGTACTGTAACTGGAAGTGCAGCCCTGAGC 540

QY 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyPheValAlaMetHisLeu 365
Db 541 GCATCGTCATCTCAGCCACTCTGGTCATCTCTGGCATCTTTTGTGGCCATGCACTG 600

QY 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyroGluLeuThrGluAsp 385
Db 601 TTGGCTAAACTGGCACCTGCAGCGCATGGAGGGGCAGATGTATGAGATCAGGAGGAC 660

QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyPro 400
Db 661 ACAGCCAGCAGTTGGCTGTGCCAACCGACGCTCTCCCTATATACCC 705

RESULT 9
ADE07179 standard; DNA; 1534 BP.
XX
AC ADE07179;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel coding sequence (useful for identifying genetic disorders) #245.
XX
KW novel gene; novel protein; tissue marker; molecular weight marker;
XX chromosome marker; genetic disorder; gene; ds.
XX Unidentified.
XX WC0203054152-A2.
XX
```

PD 03-JUL-2003.

XX 10-DEC-2002; 2002WO-US039555.

XX 10-DEC-2001; 2001US-0339739P.

PR 11-DEC-2001; 2001US-0339453P.

PR 14-MAR-2002; 2002US-0365091P.

PR 14-MAR-2002; 2002US-0365384P.

PR 12-APR-2002; 2002US-0372381P.

PR 12-APR-2002; 2002US-0372615P.

PR 22-APR-2002; 2002US-0012855P.

PR 24-APR-2002; 2002US-0376045P.

(HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;

PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RI, Wang Z;

PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

XX WPI: 2003-569235/53.

DR P-PSDB; ADE08090.

XX New polynucleotides, useful for expressing recombinant proteins for

PT analysis, characterization or therapeutic use, or as markers for tissues

PT in which the corresponding protein is preferentially expressed.

XX Claim 1; SEQ ID NO 245; 1177bp; English.

XX The invention comprises the amino acid and coding sequences of novel

CC proteins. The DNA and protein sequences of the invention are useful as:

CC markers for tissues in which the corresponding protein is preferentially

CC expressed; as molecular weight markers on gels; as chromosome markers or

CC tags; to identify chromosomes or to map related gene positions; and to

CC compare with endogenous DNA sequences in patients to identify potential

CC genetic disorders. The present DNA sequence represents a gene of the

CC invention.

XX Sequence 1534 BP; 462 A; 377 C; 355 G; 340 T; 0 U; 0 Other;

SQ

Alignment Scores:

Pred. No.: 3,3e-54 Length: 1534

Score: 987.00 Matches: 212

Percent Similarity: 62.89% Conservative: 49

Best Local Similarity: 51.08% Mismatches: 90

Query Match: 46.06% Indels: 64

DB: 9 Gaps: 12

US-10-029-020-14_COPY_1_400 (1-400) x ADE07179 (1-1534)

QY 1 MetAspValLysGluArgLysProTyArgSerLeuThrArg---ArgArgAspAlaGlu 19

Db 133 ATGGATGTGAAGAACGCGAGCTTACTGCTCCCTGACCAAGACGACGAGAGAGGAA 192

QY 20 ArgArgTyThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38

Db 193 CGGCGCTACACAAATTCCTCCGACAAATGAGGAGTCCGGGTACCCACACAGAGTCC 252

QY 39 TyrSerSerSerGluThrLeuLysAlaTyArgGlnAsp---AlaArgLeuAlaTyGly 57

Db 253 TACAGTTCCAGCGAGACATTTGAAGCTTTTGTATCATGATTCCTCGGGTGTCTTACGCG 312

QY 58 SerArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsn 77

Db 313 AACAGAGTCAGGATTTGGTTTCACAGAGAAGCAGCAGTTCCTACTAGACAAGACAGAAT 372

QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyArg 97

Db 373 TTACCTAAGCGAGTTAGGAGTTTGTGAACACGACCACTCGAGAGACTGGCATTTGT 432

QY 98 ThrAspIleGlyLeuProGlnCysGlyTySerMetGlyAlaGlySerAspAlaAspMet 117

Db 433 GCGGAATGGGCTCCCTCACAGAGGTTACTCTATCAGTCAGGGTCAGATGCTGATACT 492

CC and protozoal infections. The NOVX proteins can be used as immunogens to
 CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
 CC used in chromosome mapping, identifying individuals from minute
 CC biological samples (tissue typing), and in forensic identification of a
 CC biological sample. The present sequence encodes human NOV15b, which is
 CC located on chromosome 4

XX Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 2,42e-53 Length: 8645
 Score: 987.00 Matches: 212
 Percent Similarity: 62.8% Conservative: 49
 Best Local Similarity: 51.0% Mismatches: 90
 Query Match: 46.0% Indels: 64
 DB: Gaps: 12

US-10-029-020-14_COPY_1_400 (1-400) x ABQ82344 (1-8645)

Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
 Db 151 ATGGATGTGAAGAACGAGCGCTTACTGCTCCCTGACCAAGCAGCAGCAGAGAGAA 210
 Qy 20 ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
 Db 211 CGGCCCTACACAATTCCTCCGACACAATGAGGAGTGCCTGCTACCCACACAGAGTCC 270
 Qy 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
 Db 271 TACAGTCCAGCGAGACATTTGAAGCTTTGATCATGATTCCTCGCGCTGCTTTACGGC 330
 Qy 58 SerArgValLysAspIleValProGlnGluAlaGluLysPheCysArgThrGlyAlaAsn 77
 Db 331 AACAGATGAAGATTTGGTTACAGAGAGCAGCAGAGTTCCTAGCAGAGCAGAT 390
 Qy 78 PheThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArg 97
 Db 391 TTTACCTTAAGCAGTGTAGGAGTTTGTGAACAGCACTCGAAGAGGACTGGCAATTTGT 450
 Qy 98 ThrAspIleGluProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
 Db 451 CGCGAATGGGCTCCCTCACAGAGTTACTTATCMTGTCAGGGTTCAGATGCTGATCT 510
 Qy 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrPcylArgSerThrArg 137
 Db 511 GAAATGACAGTGTATGTCCTCCAGAGCATGCCATGAGACTTTGGGCGGGGGTCAA 570
 Qy 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
 Db 571 TCAGGCGCGAGCTCTCTGCTGTCAAGTCGTCCTCACTAGCCCTCACNCTGACAGATAG 630
 Qy 158 GluHisGluAsn-----ThrGluThrAspHisProGlyLeuGlnAsnHisAla 174
 Db 631 GAGCAGAAACAAAGTCCGACAGTGAAGAACCTGCAAGCAATCAAGCCAGTCT 690
 Qy 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisAla 194
 Db 691 ACCCTGACCCCTTGGCGCT-----TCCATAAGCAGCAGCTCTGACAGCATCAT--- 741
 Qy 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
 Db 742 CCATCCATCACTCTCTCAACAGAAAACCTCCCTGACCAATAGAGAACCAAGATCGCGCC 801
 Qy 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyAlaGln----- 230
 Db 802 CCG-----CCGGTGTCTTGGCCCGCCGAGCTGCACCAACCA 837
 Qy 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
 Db 838 CCGGAGTCCGTCCAGCTGCAGGACAGCTGGTCTTGGCAGTAAATACCACTGGAAAGC 897
 Qy 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269

Db 898 AGG----- 900
 Qy 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
 Db 901 -----CATTTCTATTCAAAACA 918
 Qy 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
 Db 919 GGAAACAGGTACACGCCACTGTCAGTACTGCAACCCCGAGGATACACAATGCGCATCTGCG 978
 Qy 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
 Db 979 TCTGTTTATTCAACACCTACTCGGCCACTACTAGAAACACCTATCAAGAAGTCTTTT 1038
 Qy 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
 Db 1039 AAATTCAGAAGTCTTCAAGTACTGTAGCTGGAATGCACTGCACTGTGCGGTAGGG 1098
 Qy 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
 Db 1099 GTCTCGGTGCTCTGGCAATATCTCTGTCTTATTTATAGCAATGATCTCTTTGCGCTC 1158
 Qy 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
 Db 1159 AACTGGCAGCTACAGCAGACTGAAATGACACATTTGAGATGGAAAGTGAATTTCTGAT 1218
 Qy 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
 Db 1219 ACC-----ATGCCAACAAACACTGTGTACCTTACCT 1248
 RESULT 11
 ID ABQ82343
 XX AC ABQ82343;
 DT 17-DEC-2002 (first entry)
 XX Human NOV15a encoding cDNA SEQ ID NO:35.
 XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
 KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
 KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
 KW antileukemic; antibacterial; protozoacide; antihelminthic; gene therapy;
 KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
 KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
 KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
 KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
 KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
 KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
 KW rheumatoid arthritis; gene; chromosome 4; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 CDS 151..8328
 FT /*tag= a
 FT /product= "NOV15a"
 FT /transl_except= (pos:1249..1251,aa:Ser)
 XX WO200262999-A2.
 XX 15-AUG-2002.
 XX 31-DEC-2001; 2001WO-US049976.
 XX 29-DEC-2000; 2000US-0258928P.
 XX 02-JAN-2001; 2001US-0259415P.
 XX 04-JAN-2001; 2001US-0259785P.
 XX 20-FEB-2001; 2001US-0269814P.
 XX 09-MAR-2001; 2001US-0279863P.
 XX 29-MAR-2001; 2001US-0279832P.
 XX 29-MAR-2001; 2001US-0279833P.

13-APR-2001; 2001US-0283889P.
 18-APR-2001; 2001US-0284447P.
 25-APR-2001; 2001US-0286683P.
 29-MAY-2001; 2001US-0294080P.
 16-AUG-2001; 2001US-0312915P.
 17-AUG-2001; 2001US-0313325P.
 17-SEP-2001; 2001US-0322699P.
 26-NOV-2001; 2001US-0333350P.
 (CURA-) CURAGEN CORP.
 XX
 XX
 XX
 PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
 PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
 PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
 PI Gunther E, Smithson G, Millet I, Macdougall JR;
 XX
 DR WPI; 2002-732706/79.
 DR P-PSDB; ABP53586.
 XX

PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
 PT associated disorders, such as cancers, neurological disorders, disorders
 PT of vesicular transport, gastrointestinal disorders, and autoimmune
 PT diseases.

PS Claim 8; Page 110-112; 444pp; English.

CC The present invention describes novel human proteins designated NOVX,
 CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
 CC cytoskeletal, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,
 CC antiadipic, antiinflammatory, antirheumatic, antiarthritic, virucide,
 CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
 CC protozoacide and antihelminthic activities, and can be used in gene
 CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease selected from NOVX-associated disorder, such as cancers
 CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
 CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
 CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
 CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
 CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
 CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
 CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
 CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
 CC and protozoal infections. The NOVX proteins can be used as immunogens to
 CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
 CC used in chromosome mapping, identifying individuals from minute
 CC biological samples (tissue typing), and in forensic identification of a
 CC biological sample. The present sequence encodes human NOV15a, which is
 CC located on chromosome 4
 XX

SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 2,43e-53 | Length: | 8675 |
| Score: | 987.00 | Matches: | 212 |
| Percent Similarity: | 62.89% | Conservative: | 49 |
| Best Local Similarity: | 51.08% | Mismatches: | 90 |
| Query Match: | 46.06% | Indels: | 64 |
| DB: | 6 | Gaps: | 12 |

US-10-029-020-14_COPY_1_400 (1-400) x ABQ82343 (1-8675)

| | | | |
|----|-----|--|-----|
| QY | 1 | MetAspValIysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu | 19 |
| DB | 151 | ATGGATGTGAAGACGCGCCTTACTGCTCCCTGACCAAGACGACGAGAGAGAA | 210 |
| QY | 20 | ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer | 38 |
| DB | 211 | CGCGGTACACAAATTCCTCCGACACATGAGAGATGCGGGTACCCACAGAAGTCC | 270 |
| QY | 39 | TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly | 57 |

| | | | |
|----|------|---|------|
| Db | 271 | TACAGTTCACGACGAGACATTCGAAAGCTTTTGAATCATGATTCCTCGCGCTGCTTACGGC | 330 |
| QY | 58 | SerArgValIysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsn | 77 |
| Db | 331 | AACAGAGTGAAGGATTTGGTTTCACAGAGACGACAGAGTTCCTACTAGACAGGACAGAT | 390 |
| QY | 78 | PheThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArg | 97 |
| Db | 391 | TTTACCTTAAGGCAAGTTAGGAGTTTGTGAACACGACCACTCGAAGAGGACTGGCATTTGT | 450 |
| QY | 98 | ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet | 117 |
| Db | 451 | CGGAAATGGGCTCCCTTCACAGAGGTACTCTATCAGTCGAGGTGAGATGCTGATACT | 510 |
| QY | 118 | GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg | 137 |
| Db | 511 | GAAATGAAGCAGTGATGTCCACAGACATGCTCCATGAGACTTTGGGCGAGGGGGTCAAA | 570 |
| QY | 138 | SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr | 157 |
| Db | 571 | TCAGGCGCAGCTCTCGCTGTCAAGTCGGTCCAACTCAGCCCTCAGCCCTGACAGATACG | 630 |
| QY | 158 | GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla | 174 |
| Db | 631 | GAGCAGCAAAACAAAGTCGACAGTGAAGTGAAGCACTGCAAGCAATCAAGGCGAGTCT | 690 |
| QY | 175 | ArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla | 194 |
| Db | 691 | ACCTGACGCGCTTGGCGCT-----TCCATAAGCAGCACTCTGCACAGCATCAT---741 | |
| QY | 195 | AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla | 214 |
| Db | 742 | CCATCCATCACTTCTCTCAACAGAACTCCCTGACCAATAGAAAGAACACAGAGTCCGCC | 801 |
| QY | 215 | ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln-----230 | |
| Db | 802 | CCG-----CCGCTGCTTTGGCCGCGAGCTGCAAAACCACA837 | |
| QY | 231 | ---GluProAlaHisAlaGlnGluAsnTrpLeuAsnSerAsnIleProLeuGluThr | 249 |
| Db | 838 | CCCGAGTCGCTCCAGCTGCAGGACAGCTGGTCTTGGCAGTAATGTACCACTGGAAGC | 897 |
| QY | 250 | ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp | 269 |
| Db | 898 | AGG-----900 | |
| QY | 270 | IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro | 289 |
| Db | 901 | -----CATTTCTATTCAAAACA918 | |
| QY | 290 | Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer | 308 |
| Db | 919 | GGAAACAGGTACACGCGCACTGTCAGTACTGTCAACCCAGGATACACAATGGCATCTGGC | 978 |
| QY | 309 | ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe | 328 |
| Db | 979 | TCGTGTTATTCCACCACTACTCGGCCACTACCTAGAAACACCCCTATCAGAAAGTCTTTT | 1038 |
| QY | 329 | AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal | 348 |
| Db | 1039 | AAATTCAGAAGTCTTCAAAGTACTGTAGCTGGAATGCACTGCTGTGTGCGGTAGGG | 1098 |
| QY | 349 | IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu | 368 |
| Db | 1099 | GTCTCGGTCTGCTGCAATACTCTCTGTCTATTATAGCAATGTCATCTCTTTGGCCTC | 1158 |
| QY | 369 | AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp | 385 |
| Db | 1159 | AACTGCGACACTCAGCAGACTGAAATGACACATTTTGAATGGAAGTGAATTTCTGAT | 1218 |
| QY | 386 | ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro400 | |
| Db | 1219 | ACC-----ATGCCACAAACACTGTGTCTATTACCT1248 | |

RESULT 12

ACC72052
ID ACC72052 standard; DNA; 9695 BP.

XX
AC ACC72052;

XX
DT 08-JUL-2003 (first entry)

XX
DE BC0205B gene #SEQ ID 81.

XX
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.

XX
OS Homo sapiens.

XX
PN WO2003029421-A2.

XX
PD 10-APR-2003.

XX
PF 02-OCT-2002; 2002WO-US031287.

XX
PR 03-OCT-2001; 2001US-0326526P.

PR 14-MAY-2002; 2002US-00144194.

XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;

XX
DR WPI; 2003-381623/36.

DR P-PSDB; ABR58318.

XX
PT New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.

PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.

XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 8,646-51 | Length: | 9695 |
| Score: | 948.00 | Matches: | 207 |
| Percent Similarity: | 64.06% | Conservative: | 55 |
| Best Local Similarity: | 50.61% | Mismatches: | 91 |
| Query Match: | 44.24% | Indels: | 56 |
| DB: | 7 | Gaps: | 13 |

US-10-029-020-14_COPY_1_400 (1-400) x ACC72052 (1-9695)

| | | | |
|----|-----|---|-----|
| QY | 1 | MetAspValLysGluArgLysProTyrArgSerLeuThrArgArg---AspAlaGlu | 19 |
| Db | 435 | ATGGATGTAAGACCGCGA---CACCGCTCTTGACACAGGAGCGCTGTGGCAAGAG | 491 |
| QY | 20 | ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer | 38 |
| Db | 492 | TGTCGTACCAAGCTCTCTCTGGACAGTGAGGACTGCGCGTGCCACACAGAAATCC | 551 |

| | | | |
|----|------|--|------|
| QY | 39 | TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer | 58 |
| Db | 552 | TACAGCTCCAGTGAGACTCTGAGGCTTATGACCATGACAGCAGGATATGGAAC | 611 |
| QY | 59 | ArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsn | 78 |
| Db | 612 | CGAGTCACAGACCTCATCCACGGGAGTCAGATGAGTTCCTAGACAGGAACCAACTTC | 671 |
| QY | 79 | ThrLeuArgGluLeuGluGluValThrProProHisGlyThrLeuTyrArgThr | 98 |
| Db | 672 | ACCTTGGCCAGACTGGGCTCTGTGAGCCCTCC---CCACACCGAGCGGCTACTGCTCC | 728 |
| QY | 99 | AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu | 118 |
| Db | 729 | GACATGGGGATCCTTCCACAGGGCTACTCCCTTAGCACAGGGTCTGACGCGGACTCCGAC | 788 |
| QY | 119 | AlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSer | 138 |
| Db | 789 | ACCGAGGGAGGATGCTCCAGAACACGCCATCAGACTGTGGGCGAGGNTAAATCC | 848 |
| QY | 139 | GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu | 158 |
| Db | 849 | AGGCGAGTTCGGGCTGTCCAGTCTGTGAAACTCGGCCCTTACCTGACTGACTGTGAC | 908 |
| QY | 159 | HisGluAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArg-----Leu | 176 |
| Db | 909 | AACGAAACAAATCAGATGATGAGAACCGCCCTCCGAAACCCACACAGCCGACTCTG | 968 |
| QY | 177 | ArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSer | 196 |
| Db | 969 | AGGCCCCCTCTCCACCC---CCTCACAAACACAGCTGTGCCATCACAC---TCGTCC | 1022 |
| QY | 197 | IleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn-----ProSer | 212 |
| Db | 1023 | GCCAACTCCCTCAACAGGAACTCAGTACCACCGGAGTCAGATCCACGCCCGGC | 1082 |
| QY | 213 | ProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluPro | 232 |
| Db | 1083 | CCAGCGCCCAATGAC-----CTGGCCACACACCA-----GAGTCC | 1118 |
| QY | 233 | AlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeu | 252 |
| Db | 1119 | GTTACGTTACGACAGCTGGGTGTAAACAGCAAGCTGCCACTGGAGACCGG----- | 1172 |
| QY | 253 | GlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGly | 272 |
| Db | 1172 | ----- | 1172 |
| QY | 273 | AlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLys---ProGlyGly | 291 |
| Db | 1173 | -----CACTTCCTCTTCAAGACCTCTCGGGG | 1199 |
| QY | 292 | ThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyr | 311 |
| Db | 1200 | AGCACACCTTGTTCAGCAGCTCTTCCCGGATACCTTTGACCTCAGGAACGGTTTAC | 1259 |
| QY | 312 | SerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeuLys | 331 |
| Db | 1260 | ACGCCCCCGCCCGCTGCTGCCAGAACTTTCTCCAGGAAGGCTTTCAGAGCTGAAG | 1319 |
| QY | 332 | LysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSerAla | 351 |
| Db | 1320 | AAGCCCTCCAAATACTGACGTGGAAATGTGTCCTCTCGCCATTCGCGGCGCCCTC | 1379 |
| QY | 352 | ThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisPhePheGlyLeuAsnTrpHis | 371 |
| Db | 1380 | CTCTTGCTATTTCGTGGGATTTTCATAGCAATGCACTGCTCGGACTCAATTGGCAA | 1439 |
| QY | 372 | LeuGlnProMetGluGlnMetTyrGluIleThrGluAspThrAlaSerSerThrPro | 391 |
| Db | 1440 | CTCCAGCTGCAGATGGGCACACCTTT-----AACATGGGATAAGGACCGCGCTACCA | 1493 |
| QY | 392 | ValProThrAspValSerLeuTyrPro | 400 |

Db 1494 GGAACGATGATGGCAACAATGCCA 1520
RESULT 13
AA514089
ID AA514089 standard; DNA; 9729 BP.
XX AA514089;
AC
XX 18-DEC-2001 (first entry)
XX
DE Human FCTR3f DNA sequence.
XX
KW Human; FCTR3f; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
KW astrocytoma; congenital neonatal alloimmune thrombocytopenia; infection;
KW neurological disorder; neurodegenerative disorders; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greenow type I; Corneal dystrophy-lattice type I;
KW Reiss-Bucklers corneal dystrophy; cytotatic; immunosuppressive;
KW antiallergic; antiasthmatic; antifertility; antinflammatory;
KW antidiabetic; prozoacide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antifertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3a; neurestin-like protein; FCTR3f.
XX
OS Homo sapiens.
XX
XX
XX
FH Key Location/Qualifiers
FT 5'UTR 1..209
FT /*tag= b
FT CDS 210..8384
FT /*tag= a
FT /*product= "Human FCTR3f"
FT 3'UTR 8385..9729
FT /*tag= c
XX
XX WO200166747-A2.
XX
XX
XX 13-SEP-2001.
XX
XX 05-MAR-2001; 2001WO-US0007160.
XX
XX 03-MAR-2000; 2000US-0186592P.
XX 03-MAR-2000; 2000US-0186718P.
XX 06-MAR-2000; 2000US-0187293P.
XX 06-MAR-2000; 2000US-0187294P.
XX 17-MAR-2000; 2000US-0190400P.
XX 07-APR-2000; 2000US-0196018P.
XX 03-JAN-2001; 2001US-0259548P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
XX Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX WPI; 2001-596837/67.
XX P-PSDB; AAU08681.
XX
XX Novel polypeptides designated as FCTR3 polypeptides, useful in detection,
XX prevention and treatment of a broad range of pathological states.
XX
XX Claim 9; Page 37-39; 215pp; English.
XX
XX The invention relates to human FCTR3 polypeptides, FCTR1-FCTR7, and the
XX nucleic acids encoding them. These sequences are useful for the treatment
XX or prevention of numerous disorders including myelogenous leukaemia,
XX carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal

CC alloimmune thrombocytopenia; neurological disorders, neurodegenerative
CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
CC -tooth neuropathy, demyelinating Gardner syndrome, familial
CC myelodysplastic syndrome, mental health conditions, immunological
CC disorders, allergy and infection, bronchial asthma, Avellino type
CC eosinophilia, lung diseases, reproductive disorders, infertility, male
CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
CC Corneal dystrophy-Greenow type I, Corneal dystrophy-lattice type I and
CC Reiss-Bucklers corneal dystrophy. This sequence represents DNA encoding
CC the FCTR3a homologue FCTR3f
XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,21e-44 Length: 9729
Score: 839.00 Matches: 195
Percent Similarity: 54.42% Conservative: 51
Best Local Similarity: 43.14% Mismatches: 76
Query Match: 39.15% Indels: 130
DB: 5 Gaps: 14

US-10-029-020-14_COPY_1_400 (1-400) x AA514089 (1-9729)
QY 1 MetAspVallysGluArgLysProTyrArgSerLeuThrArgArgG---AspAlaGlu 19
Db 210 ATGATGTAAGGACCGCGCA---CACCGCTCTTTGACGAGGACGCTGTGGCAAGAG 266
QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 267 TGTGCTTACACAAGCTCTCTCTGGACAGTGGAGACTGCCGGGTGCCACACAGAAATCC 326
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 327 TACAGCTCCAGTGAGACTCTGAAGGCTATGACCATGACGAGGATGCACTATGGAAC 386
QY 59 ArgVallysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhe 78
Db 387 CGAGTCACAGACCTCATCCACCGGAGTCAGATGAGTTCTCTAGACAAGAACCACTTC 446
QY 79 ThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThr 98
Db 447 ACCCTTCCGCAACTGGGCATCTGTGAGCCCTCC---CCACACCGAAGCGGCTACTGTCTCC 503
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 504 GATATGGGATCCTTACCAGGGCTACTCCTTAGCACAGGGTCTGACGGCGACTCCGAC 563
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSer 138
Db 564 ACCGAGGGGGATGTCCTCAGAACACGCCATCAGACTGTGGGGCAGAGGGATAAATCC 623
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 624 AGCGCAGTTCCGGCCTGTCCAGTCGTGAAATCCGGCCCTTACCCCTGACTGCTCTGAC 683
QY 159 HisGluAsnThrGluThrAspHisProGly----- 168
Db 684 AAGCAAAACAATCAGATGATGAGAAGCGTCTGCCATTCACCTACATCCTCGCTAGT 743
QY 168 ----- 168
Db 744 CTCCTCCCATCTGCTCAGCTGCTAGTCTCCATAATCTCCACCAGTGTAGTCCAGATG 803
QY 168 ----- 168
Db 804 CCATTGCTAGACAGCAACACCTCCCATCAAAATCATGGACACCAACCCCTGATGAGGAATC 863
QY 169 ----- 169
Db 864 TCCCCAATTACCTGCTCAGAGCATGCTCAGGGGCCCCAGCAAGCCTCCAGCAGTGGC 923

```
QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrProProProLeuSerHisAla 186
DB 924 CCTCGAACCACACAGCAGTCGACTCTGAGCCCTCTCCACCC---CCTCACAAAC 980
QY 187 HisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
DB 981 CACAGCTGTCCATCACAC---TCGTCCGCACTCCTCAACAGAACTCCTGACC 1037
QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
DB 1038 AATCGCGGAGTCAGATCATCCAGCCCGCCGCGCCAGCCCAATGAC---CTGGCCACC 1091
QY 223 GluProProAlaGlyClyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsn 242
DB 1092 ACACCA-----GAGTCGTTTCAGCTTCAGGACAGCTGGGTGCTAAAC 1133
QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln 262
DB 1134 AGCAAGTGCCTGAGACCCGG-----1157
QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
DB 1157 -----1157
QY 283 GlyHisPheLeuPheIys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
DB 1158 ---CACTTCTCTTCAAGACCTCTCGGGGAGACACACCTTGTTCAGCAGCTCTTCCCG 1214
QY 302 GlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArgSer 321
DB 1215 GGATACCTTTGACCTCAGGAACGGTTTACGCGCCCGCGCTGTGCTGCCAGGAAT 1274
QY 322 ThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCys 341
DB 1275 ACTTCTTCAGGAAGCTTCAAGCTGAAGAAGCCCTCAAAATCTGCAGCTGGAATGT 1334
QY 342 AlaAlaLeuSerAlaTleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal 361
DB 1335 GTGGCTCTCGCCCATTCGCGGCGCTCTCTTGGCTATTGCTGGCGATTTCATA 1394
QY 362 AlaMetHisLeuPheGlyLeuAsnTrpHisLeuGln 373
DB 1395 -----GTGCCCTGGTCTTGAAA 1412
RESULT 14
ADB32028
ID ADB32028 standard; cDNA; 9729 BP.
XX
AC ADB32028;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human FCTR3f cDNA.
XX
KW Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurofibromatosis;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW lattice corneal dystrophy.
XX
OS Homo sapiens.
XX
PN US2003087816-A1.
```

```
XX
PD
XX
PF 05-MAR-2001; 2001US-00800198.
XX
PR 03-MAR-2000; 2000US-0186592P.
PA (VERM/) VERMET C.
PA (FERN/) FERNANDES E.
PA (SHIM/) SHIMKETS R.
PA (HERR/) HERRMANN J.
PA (MAJU/) MAJUMDER K.
PA (MACD/) MACDOUGALL J.
PA (MISH/) MISHRA V.
PA (MEZE/) MEZES P.
PA (RAST/) RASTELLI L.
XX
PI Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
PI MacDougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI: 2003-625633/59.
DR P-PSDB; ADB32029.
XX
PT New FCTR3 polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTR3-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.
PS Claim 9; Page 34-37; 155pp; English.
XX
CC The invention relates to FCTR3 polypeptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTR3 polypeptide, such as colorectal cancer,
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, neurofibromatosis, familial myelodysplastic
CC neurodegenerative disorders, nerve trauma, Charcot-Marie-Tooth syndrome,
CC conditions, immunological disorders, allergy and infection, asthma, lung
CC diseases, male and female reproductive disorders, deafness, glycoprotein
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTR3 polypeptide of the invention.
XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 8,21e-44 Length: 9729
Score: 839.00 Matches: 195
Percent Similarity: 54.42% Conservatives: 51
Best Local Similarity: 43.14% Mismatches: 76
Query Match: 39.15% Indels: 130
DB: 14 Gaps: 14
US-10-029-020-14_COPY_1_400 (1-400) x ADB32028 (1-9729)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArg---AspAlaGlu 19
DB 210 ATGGATGTAAGACGACCGGGA---CACCGCTCTTTGACCAGAGGACGCTGTGCCAAGAG 256
QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
DB 267 TGTTCGCTACACAGCTCTCTCTGAGACGTGAGGACTGCCGGGTGCCACACAGAATCC 326
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
DB 327 TACAGCTCCAGTCTGAGACTCTGAAGGCTTATGACCATGACAGCAGGATGCATATGGAAC 386
```

QY 59 ArgValIysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhe 78
 Db 387 CGAGTCACAGACCTCATCCACGGGAGTCAGATGAGTTTCTGACAGGAAACCACTTC 446
 QY 79 ThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThr 98
 Db 447 ACCCTTGGCAACTGGCATCTGTGAGCCCTCC---CCACACCGAAGGGCTACTGTCTCC 503
 QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
 Db 504 GACATGGGATCCTTCACAGGGGTACTCCCTTAGCACAGGGTCTGACGCGGACTCCGAC 563
 QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSer 138
 Db 564 ACCGAGGGAGGATGTTCCAGAAACGCCATCAGACTGTGGGCGAGGGATAAATPCC 623
 QY 139 GlyArgSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
 Db 624 AGGCGCAGTTCCGGCCTGTCCAGTCGTGAAACTCGGCCCTTACCCTGACTGACTCTGAC 683
 QY 159 HisGluAsnThrGluThrAspHisProGly----- 168
 Db 684 AAGCAAAACAAATCAGATGATGAGAACGGTGTCTCCATTCACCTCATCATCTCGCCTAGT 743
 QY 168 ----- 168
 Db 744 CTCTCCCATCTGCTCAGTCGCTAGTCCATTAATCTCCACAGTTAGTGCAGATG 803
 QY 168 ----- 168
 Db 804 CCATTGCTAGACAGCAACCTCCCATCAATCATGAGACCAACCCCTGATGAGGAATTC 863
 QY 169 -----Gly 169
 Db 864 TCCCCCAATTACTACTGCTCAGAGCATGCTCAGGGCCCGCCAGCAAGCCTCCAGCAGTGC 923
 QY 170 LeuGluAsnHis-----AlaArgLeuArgThrProProProLeuSerHisAla 186
 Db 924 CTTCCGNACACACACAGCCAGTGCATCTGAGGCCCTCTCCACCC---CCTCAAC 980
 QY 187 HisThrProAsnGlnHisHisAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
 Db 981 CACACGGTGTCCCATCACAC---TCGTCCGCAACTCCCTCAACAGGAACCTCACTGACC 1037
 QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
 Db 1038 AATCGCGGAGTCAGATCCACGCCCGCCCGCCAGCCCAATGAC---CTGGCCACC 1091
 QY 223 GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsn 242
 Db 1092 ACACCA-----GAGTCGTTTCAGCTTCAGGACAGCTGGGTGCTAAAC 1133
 QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyGlnProPheLeuGlyThrLeuGln 262
 Db 1134 AGCAAGCTGCCACTGGAGACCGCG----- 1157
 QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
 Db 1157 ----- 1157
 QY 283 GlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
 Db 1158 ---CACTTCTCTTCAGACCTCTCGGGGAGACACACCTTGTTCAGCAGCTCTTCCCG 1214
 QY 302 GlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArgSer 321
 Db 1215 GGATACCTTTGACCTCAGAACGGTTTACACGCCCGCCCGCCCTGCTGCGCAGGAAT 1274
 QY 322 ThrPheAlaArgProAlaAsnLeuLysProSerLysTyrCysAsnTrpLysCys 341
 Db 1275 ACTTTCACAGGAGGCTTTTCAGAGCTGAAGAGCCCTCCAAATACCTGACGCTGGAATGT 1334
 QY 342 AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal 361

Db 1335 GCTCCCTCTCCGCCATTCGCGGCCCTCTCTTGGCTATTTGCTGGCGTATTTCATA 1394
 QY 362 AlaMetHisLeuPheGlyLeuAsnTrpHisLeuGln 373
 Db 1395 -----GTGCCCTGGTCGTTGAAA 1412
 RESULT 15
 AAS14085
 ID AAS14085 standard; DNA; 9826 BP.
 XX AAS14085;
 AC AAS14085;
 XX 18-DEC-2001 (first entry)
 XX Human FCTR3b DNA sequence.
 XX Human; FCTR3b; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
 KW astrocytoma; congenital neonatal alloimmune thrombocytopaenia; infection;
 KW neurological disorder; neurodegenerative disorders; nerve trauma;
 KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
 KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
 KW mental health condition; immunological disorder; allergy; infertility;
 KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
 KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
 KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
 KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
 KW Spino cerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
 KW Corneal dystrophy-Greenow type I; Corneal dystrophy-lattice type I;
 KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
 KW anti-allergic; antiasthmatic; antiinfertility; antiinflammatory;
 KW antidiabetic; prozoacide; hepatotropic; virucide; ophthalmological;
 KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
 KW gene therapy; FCTR3b; neurestin-like protein.
 XX Homo sapiens.
 Key Location/Qualifiers
 5'UTR 1..279
 CDS /tag= b
 280..8481
 /tag= a
 /product= "Human FCTR3b"
 3'UTR 8482..9826
 /tag= c
 WO200166747-A2.
 13-SEP-2001.
 05-MAR-2001; 2001WO-US007160.
 03-MAR-2000; 2000US-0186592P.
 03-MAR-2000; 2000US-0186718P.
 06-MAR-2000; 2000US-0187293P.
 06-MAR-2000; 2000US-0187294P.
 17-MAR-2000; 2000US-0190400P.
 07-APR-2000; 2000US-0196018P.
 03-JAN-2001; 2001US-0259548P.
 (CURA-) CURAGEN CORP.
 Vernet CAM, Fernandes E, Shinkets RA, Herrmann JL, Majumder K;
 Macdougall J, Mishra V, Mezes PS, Rastelli L;
 WPI; 2001-596837/67.
 P-PSDB; AAU08680.
 Novel polypeptides designated as FCTR3 polypeptides, useful in detection,
 prevention and treatment of a broad range of pathological states.
 Claim 9; Page 33-35; 215pp; English.

CC The invention relates to human FCTR polypeptides, FCTR1-FCTR7, and the
 CC nucleic acids encoding them. These sequences are useful for the treatment
 CC or prevention of numerous disorders including myelogenous leukaemia,
 CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
 CC autoimmune thrombocytopenia, neurological disorders, neurodegenerative
 CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
 CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
 CC myelodysplastic syndrome, mental health conditions, immunological
 CC disorders, allergy and infection, bronchial asthma, Avelino type
 CC eosinophilia, lung diseases, reproductive disorders, infertility, male
 CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
 CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
 CC disorders, pancreatic diseases such as diabetes, Schistosoma mansonii
 CC infection, spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
 CC Corneal dystrophy-Greenough type I, Corneal dystrophy-lattice type I and
 CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
 CC FCTR3b, a neurotensin-like protein
 XX
 SQ Sequence 9826 BP; 2567 A; 2643 C; 2479 G; 2137 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.3e-44 Length: 9826
 Score: 839.00 Matches: 195
 Percent Similarity: 54.42% Conservative: 51
 Best Local Similarity: 43.14% Mismatches: 76
 Query Match: 39.15% Indels: 130
 DB: 5 Gaps: 14

US-10-029-020-14_COPY_1_400 (1-400) x AAS14085 (1-9826)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArg---AspAlaGlu 19
 DB 280 AFGATGTAAGACCGCGGA---CACCGCTCTTGACACAGGAGCGCTGTGGCAAGAG 336
 QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluLysAlaPro---GlnLysSer 38
 DB 337 TGTGCTACACAAGCTCTCTGGACAGTGGAGGACTGCGGGTGGCCACACAGAATCC 396
 QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
 DB 397 TACAGCTCCAGTGAGACTCTGAAGCGCTATGACCATGACAGCAGGATGACTATGGAAC 456
 QY 59 ArgValLysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPhe 78
 DB 457 CGAGTCACAGACTCATCCACCGGAGTCAGATGAGTTCTTAGACAAGAACCACTTC 516
 QY 79 ThrLeuArgGluLeuGluGluValThrProProHisGlyThrLeuTyrArgThr 98
 DB 517 ACCCTTGCGGAACGTGGCATCTGTGAGCCCTCC---CCACACGAGCGGCTACTGCTCC 573
 QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
 DB 574 GATATGGGATCTTCCACGAGGCTACTCCCTTAGCACAGGGTCTGACGCCGACTCCGAC 633
 QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSer 138
 DB 634 ACCGAGGGAGGATGTCTCCAGAACCGCATCAGACTGTGGGGCAGAGGGATAAATCC 693
 QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
 DB 694 AGCGCAGATTCCGGCTGTCCAGTCTGTGAATACTCGGCCCTTACCCCTGACTGACTGAC 753
 QY 159 HisGluAsnThrGluThrAspHisProGly----- 168
 DB 754 AAGGAAAAAATCATCATGATGAGACGGTCTGTCCTCCATCCACTACCTCGCCTAGT 813
 QY 168 ----- 168
 DB 814 CTCTCCCATCTGCTCAGCTGCCTAGTCTCCATAATCTCCACCACTAGTCTGCCAGATG 873
 QY 168 ----- 168
 DB 874 CCATTGCTAGACAGACACCTCCCATCAATCATGACACCAACCTGATGAGGAATTC 933

QY 169 -----Gly 169
 DB 934 TCCCCCAATTACCTCTCAGACATGCTCAGGGCCCCCAGCAAGCTCCAGCAGTGGC 993
 QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrProProProProLeuSerHisAla 186
 DB 994 CCTCGGAACCAACACAGCAGCTCGACTCTGAGGCCCTCTCCACCC---CCTCACAAAC 1050
 QY 187 HisThrProAsnGlnHisAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
 DB 1051 CACACGCTGTCCCATCACAC---TCGTCCGCAACTCTCTCAACAGAACTCATGTACC 1107
 QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
 DB 1108 AATCGCGGAGTCAGATCCACGCCCGCCGAGCCCAATGAC-----CTGGCCACC 1161
 QY 223 GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuAsn 242
 DB 1162 ACACCA-----GAGTCCGTTTCAGCTTCAGGACAGCTGGGTGCTAAAC 1203
 QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln 262
 DB 1204 AGCAACGTCCTCAAGACCTCTCTCGGGAGCAGACACCTCTTTCAGCAGCTCTTCCCG 1227
 QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
 DB 1227 ----- 1227
 QY 283 GlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
 DB 1228 ---CACTTCTCTTCAAGACCTCTCTCGGGAGCAGACACCTCTTTCAGCAGCTCTTCCCG 1284
 QY 302 GlyTyrProLeuThrSerSerThrValTyrSerProProProProArgProLeuProArgSer 321
 DB 1285 GGATACCTCTTACCTCAGAAACGGTTTACACGCCCCCGCCCTGCTGCTGCCAGGAAT 1344
 QY 322 ThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCys 341
 DB 1345 ACTTCTCCAGGAGGCTTCAAGCTGAAGAAGCCCTCCCAATACTGCAGCTGGAATGT 1404
 QY 342 AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal 361
 DB 1405 GCTGCCCTCTCGCCCATTCGCCGCGCCCTCTCTGGCTATTCTTGTGCGGTATTTCATA 1464
 QY 362 AlaMetHisLeuPheGlyLeuAsnTrpHisLeuGln 373
 DB 1465 -----GTGCCCTGGTCTGTGAAA 1482

Search completed: August 14, 2004, 02:28:40

Job time : 745.508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 801.682 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_1_400

Perfect score: 2143

Sequence: 1 MDVKERKPYRSLRRRAER.....EITDTSASWPVPTDVSLLP 400

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO.spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519
-DB=Published Applications NA -OFT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MTN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020@cgn2_1_1 2156 -runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----|-------------|
|------------|-------|-------|--------|----|-------------|

ALIGNMENTS

RESULT 1

US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029236A1
; GENERAL INFORMATION:
; APPLICANT: Alsbetook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

| | 1 | 2143 | 100.0 | 8354 | 13 | US-10-383-201-43 | Sequence 43, Appl |
|----|-------|-------|-------|------|--------------------|---------------------|-------------------|
| 2 | 2143 | 100.0 | 8354 | 13 | US-10-029-020-13 | Sequence 13, Appl | |
| 3 | 2138 | 99.8 | 1680 | 15 | US-10-290-578-3 | Sequence 3, Appl | |
| 4 | 2138 | 99.8 | 3111 | 9 | US-09-773-517-12 | Sequence 12, Appl | |
| 5 | 2138 | 99.8 | 3111 | 9 | US-09-792-025-12 | Sequence 12, Appl | |
| 6 | 2138 | 99.8 | 3111 | 9 | US-09-849-868-12 | Sequence 12, Appl | |
| 7 | 2138 | 99.8 | 3111 | 15 | US-10-290-578-1 | Sequence 1, Appl | |
| 8 | 2138 | 99.8 | 3111 | 15 | US-10-453-183-12 | Sequence 12, Appl | |
| 9 | 2130 | 99.4 | 8355 | 13 | US-10-383-201-55 | Sequence 55, Appl | |
| 10 | 2098 | 97.9 | 8438 | 13 | US-10-042-865-1 | Sequence 1, Appl | |
| 11 | 1285 | 60.0 | 2387 | 15 | US-10-290-578-11 | Sequence 11, Appl | |
| 12 | 987 | 46.1 | 8645 | 17 | US-10-038-854-37 | Sequence 37, Appl | |
| 13 | 987 | 46.1 | 8675 | 17 | US-10-038-854-35 | Sequence 35, Appl | |
| 14 | 948 | 44.2 | 9695 | 16 | US-10-144-194A-81 | Sequence 81, Appl | |
| 15 | 920 | 42.9 | 8689 | 9 | US-09-808-602-78 | Sequence 78, Appl | |
| 16 | 920 | 42.9 | 8689 | 9 | US-09-800-198-66 | Sequence 66, Appl | |
| 17 | 917 | 42.8 | 2496 | 10 | US-09-808-602-75 | Sequence 75, Appl | |
| 18 | 917 | 42.8 | 2496 | 10 | US-09-800-198-63 | Sequence 63, Appl | |
| 19 | 916 | 42.7 | 8797 | 9 | US-09-808-602-74 | Sequence 74, Appl | |
| 20 | 916 | 42.7 | 8797 | 9 | US-09-808-602-77 | Sequence 77, Appl | |
| 21 | 916 | 42.7 | 8797 | 10 | US-09-800-198-62 | Sequence 62, Appl | |
| 22 | 916 | 42.7 | 8797 | 10 | US-09-800-198-65 | Sequence 65, Appl | |
| 23 | 897 | 41.9 | 8409 | 9 | US-09-808-602-79 | Sequence 79, Appl | |
| 24 | 897 | 41.9 | 8409 | 10 | US-09-800-198-67 | Sequence 67, Appl | |
| 25 | 871 | 40.6 | 8575 | 13 | US-10-072-012-143 | Sequence 143, Appl | |
| 26 | 839 | 39.2 | 9729 | 9 | US-09-808-602-12 | Sequence 12, Appl | |
| 27 | 839 | 39.2 | 9826 | 9 | US-09-808-602-7 | Sequence 7, Appl | |
| 28 | 839 | 39.2 | 9826 | 10 | US-09-800-198-7 | Sequence 7, Appl | |
| 29 | 839 | 39.2 | 9826 | 16 | US-10-295-027-927 | Sequence 927, Appl | |
| 30 | 734.5 | 34.3 | 12880 | 16 | US-10-038-854-39 | Sequence 39, Appl | |
| 31 | 653.5 | 30.5 | 8473 | 17 | US-10-038-854-41 | Sequence 41, Appl | |
| 32 | 653.5 | 30.5 | 8487 | 17 | US-09-998-966-17 | Sequence 17, Appl | |
| 33 | 534 | 24.9 | 1755 | 15 | US-10-004-415-17 | Sequence 17, Appl | |
| 34 | 534 | 24.9 | 1755 | 16 | US-10-384-974-17 | Sequence 17, Appl | |
| 35 | 534 | 24.9 | 1755 | 16 | US-10-094-749-710 | Sequence 710, Appl | |
| 36 | 508 | 23.7 | 1727 | 16 | US-10-144-194A-79 | Sequence 79, Appl | |
| 37 | 508 | 23.7 | 9058 | 16 | US-10-029-386-1356 | Sequence 1356, Appl | |
| 38 | 454 | 21.2 | 527 | 15 | US-09-808-602-5 | Sequence 5, Appl | |
| 39 | 434 | 20.3 | 1430 | 10 | US-09-800-198-5 | Sequence 5, Appl | |
| 40 | 434 | 20.3 | 1431 | 10 | US-09-977-418-21 | Sequence 21, Appl | |
| 41 | 434 | 20.3 | 1431 | 10 | US-09-977-033A-21 | Sequence 21, Appl | |
| 42 | 434 | 20.3 | 1431 | 10 | US-09-977-751C-21 | Sequence 21, Appl | |
| 43 | 434 | 20.3 | 1431 | 10 | US-09-977-639A-21 | Sequence 21, Appl | |
| 44 | 434 | 20.3 | 1431 | 11 | US-09-977-819B-21 | Sequence 21, Appl | |
| 45 | 434 | 20.3 | 1431 | 11 | | | |

| PRIORITY | | FILING DATE | | 2002-09-23 | |
|---|--|---------------|------|------------|--|
| PRIOR APPLICATION NUMBER: | 10/051,874 | | | | |
| PRIOR FILING DATE: | 2002-01-16 | | | | |
| PRIOR APPLICATION NUMBER: | 60/366,928 | | | | |
| PRIOR FILING DATE: | 2002-03-22 | | | | |
| PRIOR APPLICATION NUMBER: | 10/055,877 | | | | |
| PRIOR FILING DATE: | 2002-01-22 | | | | |
| NUMBER OF SEQ ID NOS: | 155 | | | | |
| SOFTWARE: | Curaseq1 version 0.1 | | | | |
| SEQ ID NO 43 | | | | | |
| LENGTH: | 8354 | | | | |
| TYPE: | DNA | | | | |
| ORGANISM: | Homo sapiens | | | | |
| FEATURE: | | | | | |
| NAME/KEY: | CDS | | | | |
| LOCATION: | (35)..(8341) | | | | |
| US-10-383-201-43 | | | | | |
| Alignment Scores: | | | | | |
| Pred. No.: | 4,85e-207 | Length: | 8354 | | |
| Score: | 2143.00 | Matches: | 400 | | |
| Percent Similarity: | 100.00% | Conservative: | 0 | | |
| Best Local Similarity: | 100.00% | Mismatches: | 0 | | |
| Query Match: | 100.00% | Indels: | 0 | | |
| DB: | 13 | Gaps: | 0 | | |
| US-10-029-020-14_COPY_1_400 (1-400) x US-10-383-201-43 (1-8354) | | | | | |
| QY | 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg | 20 | | | |
| DB | 35 ATGGACGTGAAGAGGAGGAGGCTTACCGCTCGCTGACCCGGCGCCGACCCGAGCCG | 94 | | | |
| QY | 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer | 40 | | | |
| DB | 95 CGCTACACAGCTCGTCCGCGACAGGAGGAGGCAAGCCCGCAGAAATCGTACAGC | 154 | | | |
| QY | 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal | 60 | | | |
| DB | 155 TCCAGCGAGACCTGAAGGCTTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 214 | | | |
| QY | 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu | 80 | | | |
| DB | 215 AAGGACATTGTCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 274 | | | |
| QY | 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle | 100 | | | |
| DB | 275 CGGAGCTGGGCTGGAGAGGTACGCCCCCTCACGGAGCCCTGTACCGGACAGACAT | 334 | | | |
| QY | 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp | 120 | | | |
| DB | 335 GGCTTGCCTCAATCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATGAGGCTGAC | 394 | | | |
| QY | 121 ThrValLeuSerProGlnHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg | 140 | | | |
| DB | 395 ACGGTGCTGTCCCTGAGCACCCCGGTCTGTGGGGGCGGAGCACCGGTGAGGCGGC | 454 | | | |
| QY | 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu | 160 | | | |
| DB | 455 AGCTCTGCTGTCCAGCGGGCCATTCCATCTCACACTCACCGACACCGGAGCATGAA | 514 | | | |
| QY | 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro | 180 | | | |
| DB | 515 AACACTGAGACTGATCATCCGGCGGCTTGCAGAACACGCGCGGTCTCGGACGCGCGC | 574 | | | |
| QY | 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu | 200 | | | |
| DB | 575 CGCGCGCTCTCGCACCGCCACACCCCGACACGACCGCGGCTTCCATTAACTCCCTG | 634 | | | |
| QY | 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu | 220 | | | |
| DB | 635 AACCGGGGCACTTTCACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 694 | | | |
| QY | 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu | 240 | | | |

Db 874 CGCGCGCTCTGACGCGCCACACCCCAACAGCAGCAGCAGCGGCTCCATTAACTCCCTG 933
Qy 201 AsnArgGlyAsnPhThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCGCGGCGCCACGACCACTCGCTC 993
Qy 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGAGAGCCCTTCCGCGGCGCGCCAGGAGCTGCCAGCGCCAGGAGAACTGGCTG 1053
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCAACAGCAACATCCCTCGAGACCAAGAACTTAGCAGCAGCACTTCTAGGAGCA 1113
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTy 280
Db 1114 TTGAGGAGCAACCTCATTTAGATGAGACATTTCTCGGCGCTCCCGCATGATGGGCTTAC 1173
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTACGGGCACTTCTCTTCAAGCTTGAGGACCTCCCGCTCTTCTGACCAACATCA 1233
Qy 301 ProGlyTyProLeuThrSerSerThrValTySerProProProArgProLeuProArg 320
Db 1234 CCAGGCTACCCACTGACCTCAGCAGCAGTGTACTCTCTCCGCGCGGACCCCTGCCCGC 1293
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyTyCysAsnTrpLys 340
Db 1294 AGCACTTCCGCGCGCGGCTTTAACTCAAGAGGAGGCTCCAACTGTAAGTCTGTAAG 1353
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyPhe 360
Db 1354 TGCCAGCCCTGAGCGCATCGTCTACAGCCACTCTGGTCATCTCTGCTGGCATACTTT 1413
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTy 380
Db 1414 GTGGCCATGCACCTGTTGGCTTAACTGGCACCTGCGACCTGCGAGCGAGTGTAT 1473
Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGTGGCTGTGCAACCGACGCTCTCCCTATAGCCCC 1533

RESULT 5

US-09-792-025-12
; Sequence 12, Application US/09792025
; Patent No. US2002042087A1

GENERAL INFORMATION:

APPLICANT: Sliwowski, Mark

Kern, Jeffrey

TITLE OF INVENTION: Use of Heregulin as a Growth Factor

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

SEQUENCING: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/792,025

FILING DATE: 23-Feb-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/020,598

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Schwartz, Timothy R.

REGISTRATION NUMBER: 32171
REFERENCE/DOCKET NUMBER: P1145

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-7467

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 3111 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

MOLECULE TYPE: nucleic acid

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-792-025-12

Alignment Scores:

Pred. No.: 4,19e-207 Length: 3111

Score: 2138.00 Matches: 399

Percent Similarity: 99.75% Conservative: 0

Best Local Similarity: 99.75% Mismatches: 1

Query Match: 99.77% Indels: 0

DB: 9 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-09-792-025-12 (1-3111)

Qy 1 MetAspValLysGluArgLysProTyArgSerLeuThrArgArgAspAlaGluArg 20

Db 334 ATGACGTGAAGAGAGAGAGCCCTTACGCTCGCTGACCGCGCGCGGACCGGAGCGC 393

Qy 21 ArgTyThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTySer 40

Db 394 CGCTACACCACTCGTCCGCGGACAGGAGAGGCAAGCCCGCAGAAATCGTACAGC 453

Qy 41 SerSerGluThrLeuLysAlaTyAspGlnAspAlaArgLeuAlaTyGlySerArgVal 60

Db 454 TCCAGCGAGACCTTGAAGGCTACGACAGACCGCCCGCTAGCTATGGCAGCGCGGTC 513

Qy 61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPhThrLeu 80

Db 514 AAGACATTGTCCGCGAGGAGCGGAGGAATCTGCCGCGACAGGTGCCAATTCACTTCC 573

Qy 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyArgThrAspIle 100

Db 574 CGGAGCTGGGCTGGAAAGTAACGCCCTCACGGGACCCCTGTACCGGACAGACATT 633

Qy 101 GlyLeuProGlnCysGlyTySerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120

Db 634 GGCTCCCCCACTGGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693

Qy 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140

Db 694 ACGGTGCTGTCCCTGAGCACCCGCTGCGTCTGTGGGGCGGAGCACACGGTCAGGGCGC 753

Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160

Db 754 AGCTCTGCTGTCCAGCGCGGCAATTCCAATCTCACACTCCCGACACCGAGCATGAA 813

Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180

Db 814 AACACTGAGACTGATCATCCGGGCGGCTGCGAAACACCGCGGCTCCGAGCGCGCGC 873

Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200

Db 874 CGCGCGCTCTCGCACGCCCAACCCCAACAGCAGCACCGCGGCTCCATTAATCCCTG 933

Qy 201 AsnArgGlyAsnPhThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220

Db 934 AACCGGGCAACTTCACGCGGAGGAGCAACCCCGAGCGCGCGCCCGGACCACTCGCTC 993

Qy 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240

Db 994 TCCGAGAGCCCCCTGCCGCGCGCGGCGGAGGCTTGCACCGCCAGGAGAACTGGCTG 1053

```
QY 241 LeuAsnSerAsnLeuProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCTCCAGACCGAGAACCTAGCAGAGAGCCATTCCTAGGAGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACCACTCATTGAGATGACATTTCTCGCGGCTCCCGGCATGATGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTCAGGGGCACTTCCTCTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCAATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProProLeuProArg 320
Db 1234 CCAGGGTACCACTGAGCTCAGCAGAGTGATCTCTCCCGCCCGACCCCTGCCCCG 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCGCGCGCGCTTTAACTCAAGAGCCCTCCCAAGTACTGTAACCTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TCGCAGAGCCCTGAGCGCCATCTCATCTCAGCCACTCTGGTCTATCTCTGTCGCACTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACTGTTGGCCCTAACTGCGCACTGCGAGCCGATGGAGGGGAGATGAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCAGTTGGCTGTGCCAACCGACGCTCTCCCTATACCCC 1533

RESULT 6
US-09-849-868-12
; Sequence 12, Application US/09849868
; Patent No. US20020081299A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: HAIR CELL DISORDERS
; FILE REFERENCE: GENENT.035VPC
; CURRENT APPLICATION NUMBER: US/09/849,868
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (334)...(2637)
US-09-849-868-12

Alignment Scores:
Pred. No.: 4, 19e-207 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservatives: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-09-849-868-12 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGCAGCTGAAGAGAGAGGAGCCCTTACCGCTCGCTGACCGCGCGCGCGAGCGCGC 393
QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACAGCTCGTCCGCGGAGCAGGAGGGGCAAGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
```

RESULT 7

US-10-290-578-1

```
Db 454 TCCAGCGAGACCCCTGAAGAGCCCTACGACGAGACCCCGCCTAGCTATGGCAGCGCGCTC 513
QY 61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 514 AAGGACATTTGCGCGAGAGGCGCGAGAAATTCGCGCACAGGTGCCAATTCACCCCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGGAGCTGGGCTGGAGAAAGTAAAGCCCTCACGGGACCCCTGTACCGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCGCCACTGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTTGAGACACCCCGTGGTCTGTGGGGCCGAGCACAGGTTCAGGGCG 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTGCTGCTCCAGCGCGGCCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGCGCGCGCTGCAAGAACACGCGCGGCTCCGGACCGCGCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaSerIleAsnSerLeu 200
Db 874 CGCGCGCTCTCGACGCCACACCCCGCACAGCACCGCGCGCTCCATTAATCCCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGGCACTTCACGCGCGAGGAGCAACCCGCGCGCGCCCGCACGACCATCTCGCTC 993
QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGAGAGAGCCCTCGCGCGCGCGCTGCGAGAGCTGCGCCACGCGCGAGGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCTCCGAGACCGAGAACCTAGGCAAGCAGCATTCCTTAGGAGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACCACTCATTGAGATGACATTTCTCGCGGCTCCCGCATGATGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTCAGGGGCACTTCCTCTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCAATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProProLeuProArg 320
Db 1234 CCAGGGTACCACTGAGCTGAGCAGCAGTGTACTCTCCCGCGCGCGACCCCTGCGCGCG 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCGCGCGCGCTTTAACTCAAGAGCCCTCCCAAGTACTGTAACCTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TCGCAGAGCCCTGAGCGCCATCTCATCTCAGCCACTCTGGTCTATCTCTGTCGCACTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACTGTTGGCCCTAACTGCGCACTGCGAGCCGATGGAGGGGAGATGAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGTGGCTGTGCCAACCGACGCTCTCCCTATACCCC 1533
```



```

; PRIOR APPLICATION NUMBER: US/09/243,198
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,866
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-453-183-12

```

| | | |
|------------------------|-----------|---------------|
| Alignment Scores: | | |
| Pred. No.: | 4.19e-207 | Length: |
| Score: | 2138.00 | Matches: |
| Percent Similarity: | 99.75% | Conservative: |
| Best Local Similarity: | 99.75% | Mismatches: |
| Query Match: | 99.77% | Indels: |
| DB: | 15 | Gaps: |
| | 0 | |

US-10-029-020-14 COPY 1 400 (1-400) x US-10-453-183-12 (1-3111)

1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
334 ATGGACGTGAGGAGAGAAACCTTACCGCTCGCTGACCGCGCGCGCGACGCGCGACGC 393

21 ArgTyrThrSerSerAlaAspSerGluGluLysAlaProGlnLysSerTyrSer 40
394 CGCTACACCACTCGTCCGCGCAGCAGGAGGGCAAGCCCGCAGAAATCGTACACG 453

Qy 41 SerSerGluThrLeuLysAlaIaTrpAspGlnAspAlaArgLeuAlaIaTrpGlySerArgVal 60
Db 454 TCCAGCGAGACCTGAGGCCCTAGACCCAGGACCGCCCGCTAGCCTATGCCAGCGCGTC 513

61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80
514 AAGGACATTGTCCCGACGAGGGCCGAGGAATTCGCGCACAGGTGCCAACTTCACCCCTG 573

81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
574 CGGGAGCTGGGGCTGGAAAGAAAGTAAACGCCGCCCTCACGGACCCCTGTACCGACACACATT 633

101 GlyLeuProGlnCysGlyTyrSerMetClyAlaGlySerAspAlaAspMetGluAlaAsp 120
 634 GGCCCTCCCCACTCGGGTACTTCATCGGGGGTGGCTCTGATGTCGCACATGGAGCTCAC 693

121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
 694 ACGTGAGTGTCCCTGAGCACCCGTCGCTCTGGAGCCGAGACACACAGGTCACGGCCG 753

141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
 754 AGTCCTCGCCTCTCCAGTCGGGCAATTCACATCTCACATCCCGACACCGACCATGAA 813

QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrPro 180
 |||||
 Db 814 AACACTGAGACTGATCATCGGCGGACTGCAGAACACCGCGGGTCCGACCGCCGCC 873

181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
|||||
874 CCGCGCTCTCGACGCGCCACACCCCCAACACGACACACGCGCGCTCTCAATTAACTCCCTG 912

201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
934 AACGGGGCAACTTTCACGCCAGCAGCAACCCAGCCGGCCCCCAAGGACGACTCGTC 993

[illegible]

241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
CTCAACAGCAACATCTCCCTCCAGACAGAACCTCCGCGACGACCTTCCTGCGGCG
1054

| | | | |
|----|------|--|------|
| Qy | 261 | LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr | 280 |
| Db | 1114 | TTGAGGACACACTCATTTGAGATGGACAATTCGGCGCCTCCGCCATGATGGGCTTAC | 1173 |
| Qy | 281 | SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer | 300 |
| Db | 1174 | AGTGACGGGCACTTCCTCTTCAAGCCCTGGAGGCACTCCCGCTCTTCTGCAACCATCA | 1233 |
| Qy | 301 | ProGlyTyrProLeuThrSerSerThrValTyrSerProProCysProLeuProArg | 320 |
| Db | 1234 | CCAGGGTACCCACTGACGTCCAGACACAGTGTACTCTCCTCGCCCCGACCCCTCGCCCGC | 1293 |
| Qy | 321 | SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys | 340 |
| Db | 1294 | AGCACCTTCGCGCGCGCGCTTTAACTCAAGAAAGCCCTCAAGTACTGTAACTGGAAG | 1353 |
| Qy | 341 | CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe | 360 |
| Db | 1354 | TGCGCAGCCCTGAGCGGCATCGTCATCTCAGCCACTCTGGTCATCTCTGCTGGCATCTTT | 1413 |
| Qy | 361 | ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr | 380 |
| Db | 1414 | GTGGCATGACACCTGTTGGCTAAACTGGCACCTCGAGCCGATGGAGGGCAGATGTAT | 1473 |
| Qy | 381 | GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro | 400 |
| Db | 1474 | GAGATCAGGAGACACAGCAGCAGTTGGCTGTGCAACCGACGCTCTCCCTATACCCC | 1533 |

RESULTS

```

US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US2004002926A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTI
; FILE REFERENCE: 21402-569A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

```

Alignment Scores:

| | | | |
|---------------------|-----------|---------------|------|
| Pred. No.: | 1.02e-205 | Length: | 8355 |
| Score: | 2130.00 | Matches: | 398 |
| Percent Similarity: | 99.50% | Conservative: | 0 |

Best Local Similarity: 99.50% Mismatches: 2
 Query Match: 99.39% Indels: 0
 DB: 13 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-10-383-201-55 (1-8355)

Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
 Db 1 ATGACCTGAGAGAGAGACCTTACCGTCGCTGACCCCGCGCCGCGAGCGC 60

Qy 21 ArgTyrThrSerSerAlaAspSerGluGluLysAlaProGlnLysSerTyrSer 40
 Db 61 CGCTACACCAAGCTCGTCGCGGACAGCGAGGAGGAGCCCGGAGAAATGCTACAGC 120

Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
 Db 121 TCCAGCGAGACCTTGAAGGCGCTACGACAGGACCGCCGCTAGCCTATGCGAGCGCGTC 180

Qy 61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80
 Db 181 AAGGACATTGTCGCGAGAGGCGCGAGAAATTCGCGGACAGGTGCACCTTCACCGCTG 240

Qy 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
 Db 241 CGGAGCTGGGGCTGGAAGAAGTAACGCCCGCTCACGGGACCTGTACCGGACAGACATT 300

Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
 Db 301 GGCTTCCCCACCTGCGGCTACTCATGGGGCTGGCTCTGATGCGGACATGGAGGCTGAC 360

Qy 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgGlyArg 140
 Db 361 ACGGTGCTGTCCTGAGACCCCGTGTGTCGTGGGGCGGAGCACCGTCAGGGCGC 420

Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
 Db 421 AGCTCTGCTGCTGCGAGCGGGCAATTCACATCTCACCTCACCGGACACCGGAGCATGAA 480

Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
 Db 481 AACACTGAGACTGATCATCGGGCGGCTGCAGAACCCAGCGGGCTCCGGAGCGCGCGC 540

Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
 Db 541 CGCGCGCTCTCGACGCCACACCCCAACAGCACCGCGGCTCCATTAATCTCCCTG 600

Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
 Db 601 AACCGGGCAACTTCAGCGGAGGAGCAACCCCGCGGCGGCGGCGGCGGCGGCGGCTC 660

Qy 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnThrLeu 240
 Db 661 TCGGAGAGCCCCCTGCGCGGCGGCGGCGGAGGCGCTGCGCCAGCGGAGGAACTGGCTG 720

Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
 Db 721 CTCAACAGCAACATCCCTGAGAGACAGAACTAGGCAAGAGCAATCTTAGGAGACA 780

Qy 261 LeuGlnAspAsnLeuLeuMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
 Db 781 TTGAGGAGCAACCTCATGAGATGGACATTCGCGGCGCTCCCGGCGGCTGATGGGCTTAC 840

Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
 Db 841 AGTGAGCGGCGACTTCTCTTCAAGCTGGAGGACCTCCCGGCTCTTCTGACCACTCA 900

Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
 Db 901 CCAGGGTACCACTGAGCTCCAGCACAGTGATCTCTCCGCGGCGGCGGCGGCGGCGG 960

Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnThrLys 340
 Db 961 AGCACCTTGGCTGGCGGCGCTTTAACTTCAAGAGCGCTCCAAAGTACTGTAAGTGAAG 1020

Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
 Db 1021 TGGCAGCCCTGAGCGCATCTCTAGCCACTCTGGTCACTCTGCTGGCATACCTT 1080

Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTyrHisLeuGlnProMetGluGlyGlnMetTyr 380
 Db 1081 GTGCCATGACCTGTGTGGCTAACTGGACCTGCGCGGATGGAGGGGCGAGATGAT 1140

Qy 381 GluIleThrGluAspThrAlaSerSerTyrProValProThrAspValSerLeuTyrPro 400
 Db 1141 GAGATCAGGAGGAGACACAGCAGCAGTGGCTGTGCGACCGGACGCTCTCCCTATACCC 1200

RESULT 10
 US-10-042-865-1
 ; Sequence 1, Application US/10042865
 ; Publication No. US20040029216A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Li, Li
 ; APPLICANT: Zerhusen, Bryan D
 ; APPLICANT: Casman, Stacie J
 ; APPLICANT: Shenoy, Suresh G
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Zhong, Mei
 ; APPLICANT: Gangolli, Esha A
 ; APPLICANT: Burgess, Catherine E
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Vernet, Corine A.M
 ; APPLICANT: Taylor, Sarah
 ; APPLICANT: Tchernev, Velizar T
 ; APPLICANT: Miller, Charles E
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Boldog, Ference L
 ; APPLICANT: Grosse, William M
 ; APPLICANT: Alsbrook II, John P
 ; APPLICANT: Gerlach, Valerie L
 ; APPLICANT: Edinger, Shlomit R
 ; APPLICANT: Rothenberg, Mark E
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: MacDougall, John
 ; APPLICANT: Malyankar, Uriel M
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Peyman, John
 ; APPLICANT: Smithson, Glennda
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Stone, David
 ; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
 ; FILE REFERENCE: 21402-537
 ; CURRENT APPLICATION NUMBER: US/10/042,865
 ; CURRENT FILING DATE: 2002-05-17
 ; PRIOR APPLICATION NUMBER: 60/260,417
 ; PRIOR FILING DATE: 2001-01-09
 ; PRIOR APPLICATION NUMBER: 60/260,831
 ; PRIOR FILING DATE: 2001-01-10
 ; PRIOR APPLICATION NUMBER: 60/272,338
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 60/274,876
 ; PRIOR FILING DATE: 2001-03-09
 ; PRIOR APPLICATION NUMBER: 60/284,704
 ; PRIOR FILING DATE: 2001-04-18
 ; NUMBER OF SEQ ID NOS: 264
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 8438
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-042-865-1

Alignment Scores: 1.86e-202 Length: 8438
 Pred. No.: 2098.00 Matches: 398
 Score:

Percent Similarity: 98.76% Conservative: 0
Best Local Similarity: 98.76% Mismatches: 2
Query Match: 97.90% Indels: 3
DB: 13 Gaps: 1

US-10-029-020-14_COPY_1_400 (1-400) x US-10-042-865-1 (1-8438)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
DB 4 ATGGACGGTGAAGGAGAGAGCCTTACCGCTCGCTGACCCGCGCGCGAGCCGAGCGC 63
QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
DB 64 CGCTACACCACTCTCTCGCGGACAGCGAGGAGGCGAAGCCCCCGACGAATCGTACAGC 123
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
DB 124 TCCAGCGAGACCTTGAAGGCTTACGACGAGGAGCGCGCTAGCCCTATGGCAGCGCGCTC 183
QY 61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80
DB 184 AAGGACATTTGCTCCCGAGAGGCCGAGGAATTTCTCGCGACAGGTGCGCACTTTCACCTG 243
QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
DB 244 CGGAGCTGGGGCTGGAGAGTACGCCCCCTCACGGGACCTGTACCGGACAGACATT 303
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
DB 304 GGCCTCCCGCCACTCGCGGTACTCCATGGGGGTGCTCTGTATGCCGACATGGAGGCTGAC 363
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
DB 364 ACGGTGCTGCTCCCTTGAGCACCCCGTGGCTCTGTGGGGCGGAGACACCGCTCAGGGCGC 423
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
DB 424 AGCTCTGCTGCTCCAGCGGGCCAAATTCATCTCACACTCACCGACACCGAGCATGAA 483
QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
DB 484 AACACTGAGACTGATCATCGGGCGGCTGCAAGACACCGCGGCTCCGGACCGCGCGC 543
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
DB 544 CGCGCGCTCTCGACGCCGCACACCCCAACAGCACACCGCGGCTCCATTAATCTCCCTG 603
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
DB 604 AACCGGGCAACTTCAACCGGAGGAGCAACCCCGCGCGCCCGCCCGACGACCACTCGCTC 663
QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnThrLeu 240
DB 664 TCCGAGAGCCCCCTGCGCGCGCGCGCCAGGAGCTGCGCCAGCCCGAGGAGAACTGGCTG 723
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
DB 724 CTCACAGCAACATCCCTCTGGAGACAGAACTTAGCAAGCAGCACTCTCTAGGACA 783
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
DB 784 TTGAGGACAACCTCATTTAGATGGACATTTCTCGGCGCTCCCGCCCATGATGGGGCTTAC 843
QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
DB 844 AGTGACGGGACATTTCTCTTCAAGCTTGGAGGACCTCCCGCTCTTTCTGCACCATCA 903
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
DB 904 CAGGGTACCACTGACGTCACAGAGTACTCTCTCCGCGCCCGACCCCTGCGCCCGC 963
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnThrLys 340

DB 964 AGCACCTTCGCTGGCGGCGCTTTAACTCAAGAGCCCTCAAGTACTGTAACTGGAG 1023
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
DB 1024 TGGCAGCCCTGAGGCGCATCTCATCCAGCACTCTGGTCACTCTGCTGGCATCTT 1083
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet--- 379
DB 1084 GTGGCCATGCACCTGTTTGGCTTAACTGGCACCTCGAGCGGATGGAGGGGCGATGAG 1143
QY 380 ----TyrGluLeuThrGluAspThrAlaSerSerTrpProValProThrAspValSerLe 398
DB 1144 GATTATGAGATCAGGAGGACACAGCCAGGTTGGCTGTGCCAACCGAGCTCTCCCT 1203
QY 398 uTyrPro 400
DB 1204 ATACCCC 1210

RESULT 11
US-10-290-578-11
: Sequence 11, Application US/10290578
: Publication No. US20030078389A1
: GENERAL INFORMATION:
: APPLICANT: Schaefer, Gabriele M.
: TITLE OF INVENTION: Gamma-Herregulin
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESS: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/290,578
FILING DATE: 08-Nov-02 US20030078389A1-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/891,845
FILING DATE: <Unknown>
APPLICATION NUMBER: 60/021640
FILING DATE: 12-Jul-96

ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P1043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881

INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2387 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-290-578-11

Alignment Scores:
Pred. No.: 1,77e-120 Length: 2387
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.96% Indels: 0
DB: 15 Gaps: 0

```
US-10-029-020-14_COPY_1_400 (1-400) x US-10-290-578-11 (1-2387)
QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCGCGCGCGCTGCAGAACACACGCGCGCTCCGGAGCGCGCGCGCGCTCTCGCAC 60
QY 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCACACCCCCAACACAGACACACGCGCGCTCATTAACCTCCGTGAACCGGGCAACTTC 120
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACCCGAGAGAGAACCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 180
QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
Db 181 GCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 240
QY 246 ProLeuGluThrArgAsnLeuGlyGlnProPheLeuGlyThrLeuGlnAsnLeu 265
Db 241 CCCTCGAGACCAAGAACCTTAGCAAGACGACCATTCCTAGGACATTCGAGGACACCTC 300
QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGACATTCCTCGCGCGCTCCCGCATGATGGGCTTACAGTGACGGGCACTTC 360
QY 286 LeuPheLysProGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 361 CTCTCAAGCTCGAGCGACCTCCCGCTCTCTGCACACCATCACCAGGGTACCCACTG 420
QY 306 ThrSerSerThrValTyrSerProProProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 ACCTCCAGCACATGTACTCTCTCCCGCGCGACCTCCCGCGCGCGCGCGCGCGCG 480
QY 326 ProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer 345
Db 481 CCGGCTTTAACTCAAGAGCGCTCCAGTACTGTAACTGGAAGTGGCGAGCGCTGAGC 540
QY 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365
Db 541 GCCATCGTCATCTCAGCACCTCTGTGTCATCTCTGTGTCATCTCTGTGTCATCTCT 600
QY 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlnMetTyrGluIleThrGluAsp 385
Db 601 TTTGGCTAAACTGGCACCTGCAGCGCATGGAGGGGCGAGATGTATGAGATCAACGGAGC 660
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCCAGCAGTTGGCTGTGCCAACCGACGCTCTCCCTATACCC 705

RESULT 12
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaofia S
; APPLICANT: Shenov, Suresh G
; APPLICANT: Rastelli, Luca

; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Macdougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 8645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-37

Alignment Scores:
Pred. No.: 2,04e-89 Length: 8645
Score: 987.00 Matches: 212
Percent Similarity: 62.89% Conservative: 49
Best Local Similarity: 51.08% Mismatches: 90
Query Match: 46.06% Indels: 64
DB: 17 Gaps: 12

US-10-029-020-14_COPY_1_400 (1-400) x US-10-038-854-37 (1-8645)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGATGTGAAGAAGACGACGCGCTTACTGCTCCCTACCAAGACGACGACGAGAAGAA 210
QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGGCGCTACACAAATTCCTCCGACAGCAATGAGGAGTGCCGGGTACCCACACAGAAGTCC 270
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTTCCAGCAGACATTTGAAAGCTTTTGATCATGATTCCTCGCGGCTGCTTACGGC 330
QY 58 SerArgValLysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsn 77
Db 331 AACAGAGTGAAGGATTTGGTTTCACAGAGAAGCAGACGAGTTCACAGCAAGGACAGAT 390
QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
Db 391 TTTACCCCTAAGCAGTTAGGAGTTTGTGAACACACGAACTCGAAGAGACTGGCATTTTGT 450
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
```

```

Db      451 GCGAAATGGGCTCCCTCACAGAGGTACTCTATCATGTCAGGGTTCAGATGCTGATACT 510
Qy      118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db      511 GAAATGAGAGAGTGTATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
Qy      138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db      571 TCAGGCGCGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
Qy      158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db      631 GAGCAGCAAAACAGTCCGACAGTGAATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690
Qy      175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisAla 194
Db      691 ACCCTGCGAGCCCTTGGCGCT-----TCCCAAGAGCAGCATCTGCGACAGCATCAT--- 741
Qy      195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db      742 CCATCCATCCTCTCTCAACAGAACTCCCTGACCAATAGAGGAACCCAGAGTCCGGCC 801
Qy      215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db      802 CGG-----CGGGTGTCTTTGCCCGCGAGCTGCAAAACCACA 837
Qy      231 ---GluProAlaHisAlaGlnGluAsnTrpLeuAsnSerAsnIleProLeuGluThr 249
Db      838 CCGAGTCCGCGAGCTGCGAGCAGCTGGGCTTGGCAGTAATGTACCATGGAAGC 897
Qy      250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db      898 AGG----- 900
Qy      270 IleLeuGlyAlaSerArgHisAspGlyAlaTrpSerAspGlyHisPheLeuPheLysPro 289
Db      901 -----CATTCCTATTCAAAACA 918
Qy      290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyTrProLeuThrSerSer 308
Db      919 GGAACAGGTACACGCCACTGTTAGTACTGCAACCCAGGATACAAATGGCATCTGGC 978
Qy      309 ThrValTySerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db      979 TCTGTTTATTACCACTACTCGGCCACTACCTAGAAACACCCCTATCAAGAAGTCTTT 1038
Qy      329 AsnLeuLysLysProSerLysTyTrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db      1039 AAATTCAGAAGTCTCTCAAGTACTGTAGTGAATGCACTGCACCTGTGTCGCGTAGGG 1098
Qy      349 IleSerAlaThrLeuValIleLeuAlaTrpPheValAlaMetHisLeuPheGlyLeu 368
Db      1099 GTCTCGGTGCTCTGGCATACTCTCTGTTATTTATAGCAATGCATCTCTTGGCCTC 1158
Qy      369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyTrGlu-----IleThrGluAsp 385
Db      1159 AACTGGCAGTACAGCAGACTGAAATGACACATTGAGATGGAAGTGAATGATCTGAT 1218
Qy      386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyTrPro 400
Db      1219 ACC-----ATGCCAACAACTGTGTCTATTACCT 1248

```

RESULT 13

```

US-10-038-854-35
; Sequence 35, Application US/10038854
; Publication No. US2004002781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J

```

```

; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shinkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderina, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shency, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2003-01-22
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35

```

```

Alignment Scores:
Pred. No.: 2,05e-89 Length: 8675
Score: 987.00 Matches: 212
Percent Similarity: 62.89% Conservative: 49
Best Local Similarity: 51.08% Mismatches: 90
Query Match: 46.06% Indels: 64
DB: 17 Gaps: 12

```

US-10-029-020-14_COPY_1_400 (1-400) x US-10-038-854-35 (1-8675)

```

Qy      1 MetAspValLysGluArgLysProTyArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db      151 ATGGATGTGAAGAACCGAGGCTTACTGCTCCCTGACCAAGAGCAGACGAGAGAGAA 210
Qy      20 ArgArgTyTrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db      211 CGGCGCTACACAATTCCTCCGACACAAATGAGGAGTCCCGGGTATCCACACAGAGTCC 270

```

```
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTTCAGCGCAGACATTGAAAGCTTTTGATCATGATTCCTCGCGGCTGCTTTACGGC 330
QY 58 SerArgValLysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsn 77
Db 331 AACAGAGTGAAGATTGGTTTCAGAGAGCAGACGAGTTCATAGACAGACAGAT 390
QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
Db 391 TTTACCTTAAGGCAGTTAGGAGTTTGTGAACCAAGCACTCGAAGAGAGCTGGCATTTTGT 450
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGGAAATGGGCTCCCTCACAGAGGTTACTTATCAGTCAGGGTCAGATGCTGATCT 510
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuThrPheGlyArgSerThrArg 137
Db 511 GAAATGACAGCAGTGATGTCCTCCAGAGCATGCGCATGAGACTTTGGGCGCAGGGGTCAAA 570
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCGCAGCTCTCGCTGTCAAGTCGTCCAACCTCAGCCCTCACCCCTGACAGATACG 630
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCAGCAAAACAAAGTCCGACAGTGAATGAGCAACCTGCAAGCAATCAAGCCAGCTCT 690
QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisAla 194
Db 691 ACCCTGAGCCCTTGGCGCT-----TCCATAAGCAGCAGCTCTGACAGCATCAT--- 741
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCATCACTCTCTCAACAGAACTCCCTGACCAATGAAGAACACAGAGTCCGGCC 801
QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CCGGCTGCTTTGCCCGCGAGCTGCAAAACCACA 837
QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCGGAGTCCCTCGAGTGGAGCAGCTGGTGGTCTTGGCGATATGATACCACTGGAAGC 897
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 898 AGG----- 900
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTTCTTATTCAAAACA 918
QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGTACAGCCACTGTTCACTACTCAACCCAGGATACACATGGCATCTGGC 978
QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTCAACCACTACTCGGCCACTACCTAGAAACACCCCTATCAAGAAGTGTCTT 1038
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAGAGAGTCTTCAAGATCTGTAGCTGGAAATGCATGCACTGTGTGCGCGTAGGG 1098
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCCGGCAATACTCTGTCTTATTTATAGCAATGATCATCTTTGGCCTC 1158
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGCTACAGCAGCTGAAATGACACATTTGAGAAATGGAAGTGAATCTCTGAT 1218
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
```

```
Db 1219 ACC-----ATGCCACAAACACTGTGTCTTACCT 1248
RESULT 14
US-10-144-194A-81
; Sequence 81, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 81
; LENGTH: 9695
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (435)..(8336)
US-10-144-194A-81
Alignment Scores:
Pred. No.: 2,2e-85 Length: 9695
Score: 948.00 Matches: 207
Percent Similarity: 64.06% Conservative: 55
Best Local Similarity: 50.61% Mismatches: 91
Query Match: 44.24% Indels: 56
DB: 13 Gaps: 13
US-10-029-020-14_COPY_1_400 (1-400) x US-10-144-194A-81 (1-9695)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArg---AspAlaGlu 19
Db 435 ATGGATGTAAGGACCGCGGA---CACCGCTCTTTGACCAGAGAGACGCTGTGGCAAGAG 491
QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 492 TGTGCGTACAAAGCTCTCTCTGGCAGTGGAGACTCGCGCTGCCACACAGAAATCC 551
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 552 TACAGTCCAGTGAGACTCTGAAGGCTATGACCATGACAGCAGGATGCACTATGGAAC 611
QY 59 ArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhe 78
Db 612 CGAGTCACAGACCTCATCCACCGGAGTCAGATGAGTTTCTAGACAAGGAACCACTTC 671
QY 79 ThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThr 98
Db 672 ACCCTTCCGAACTGGGCATCTGTAGCCCTCC---CCACACCGAAGCGGCTACTGTCTCC 728
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 729 GACATGGGATCCTTACACAGGCTACTCCTTAGCAGAGGCTGACGCGGACTCCGAC 788
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuThrPheGlyArgSerThrArgSer 138
Db 789 ACCGAGGGAGGGATGTCTCCAGAAACAGCCATCAGACTGTGGGCGCAGAGGATAAAATCC 848
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 849 AGGCGCAGTTTCCGCGCTGTCCAGTCGCTGGAATACTCGGCCCTTACCTGACTGACTCTGAC 908
QY 159 HisGluAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArg-----Leu 176
Db 909 AACGAAACAAATCAGATGATGAGACCGGCTCCCGAACCAACACACAGCCAGTCGACTCTG 968
QY 177 ArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSer 196
Db 969 AGGCCCCCTCTCCACCC---CCTCACACACAGCAGTGTCCCATCACCAC---TCGTCC 1022
```

```
QY 197 IleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn-----ProSer 212
Db 1023 GCCAACTCCCTCAACAGGAACCTCACTGACCAATCGGCGAGTCAGATCCACGCCCGGCC 1082
QY 213 ProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluPro 232
Db 1083 CCAGCGCCCAATGAC-----CTGGCCACCACACCA-----GAGTCC 1118
QY 233 AlaHisAlaGlnGluAsnTrpLeuAsnSerAsnIleProLeuGluThrArgAsnLeu 252
Db 1119 GTTCAGCTTCAGACAGCTGGGTGCTAAACAGACAGCTGCCACTGGAGACCCGG----- 1172
QY 253 GlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGly 272
Db 1172 ----- 1172
QY 273 AlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLys---ProGlyGly 291
Db 1173 -----CACTTCCTCTTCAAGACCTCTCCGGGG 1199
QY 292 ThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyr 311
Db 1200 AGCACACCTTGTTTACGAGCTCTTCCCGGATACCCCTTGACCTCAGGAACGGTTTAC 1259
QY 312 SerProProArgProLeuProLeuProArgSerThrPheAlaArgProAlaPheAsnLeuLys 331
Db 1260 ACGCCCGCCCGCCGCTGCTGCCAGGAATACTTCTCCAGGAGGCTTTCAGGCTGAAG 1319
QY 332 LysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSerAla 351
Db 1320 AAGCCCTCCAAATATCGCAGTGGAAATGTGTCGCCCTCTCCGCCATTGCCCGGCCCTC 1379
QY 352 ThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrpHis 371
Db 1380 CTCITGGCTATTTTGGTGGCTATTTTCATAGCAATGCTCTCGACTCAATTGGCA 1439
QY 372 LeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSerSerTrpPro 391
Db 1440 CTCAGCCTGCAGATGGGCACACCTTT-----AACATGGGATAAGGACCGGCTTACCA 1493
QY 392 ValProThrAspValSerLeuTyrPro 400
Db 1494 GAAACAGATGATGTGGCAACATGCCA 1520
```

RESULT 15

```
US-09-808-602-78
; Sequence 78, Application US/0908602
; Patent No. US2002015515A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US2002015515A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 78
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-808-602-78
```

```
Alignment Scores:
Pred. No.: 1,348-82 Length: 8689
Score: 920.00 Matches: 211
Percent Similarity: 54.91% Conservative: 52
Best Local Similarity: 44.05% Mismatches: 90
Query Match: 42.93% Indels: 126
DB: 9 Gaps: 14
```

US-10-029-020-14_COPY_1_400 (1-400) x US-09-808-602-78 (1-8689)

```
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArg-----AspAlaGlu 19
Db 133 ATGGATGTGAAGGATCGCGA---CATGCTCTTTGACCAGCGGACGGTGTGCCAAGGAG 189
QY 20 ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 190 TGTCTCTACACAGCTCTCTCTGACAGATGAGGACTGCCGTGTGCCACGAGAGATGCC 249
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 250 TACAGTTCCAGTGAGACCTCGAAGCTTATGACCATGACAGCAGGAATGCACCTATGGAAC 309
QY 59 ArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhe 78
Db 310 CGAGTCCAGACACTGCTGTCACCGGAGTCCGATGAGTTTCTAGACAAGGGCTAATTTTC 369
QY 79 ThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThr 98
Db 370 ACCCTGGCAGAAATGGGAATCTGCGAGCCCTCC---CCACACCGAAGTGTGTACTGTTC 426
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 427 GACATGGGGATCTCTCCAGGGCTACTCCCTGAGCAGTGGTCTGTATGCGGACTCGGAC 486
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSer 138
Db 487 ACCGAGGGAGGATGCTCTCCAGAACATGCATCAGACTGTGGGACGAGGGATAAATCG 546
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 547 AGGCGCAGCTCTGGCTTGTCCAGCGCGAGAACTCAGCCCTTACTCTGACTGATTTCTGAC 606
QY 159 HisGluAsnThrGluThrAspHisProGly----- 168
Db 607 AATGAAATAATCGGATGACGACAAATGTCGACCCATTCCACCTACATCCTCTCTAGC 666
QY 168 ----- 168
Db 667 CTCCTCCCATCTGCTCAGCTGCTAGTCCCATATCTCCACCAAGTTAGCTGCCAGATG 726
QY 168 ----- 168
Db 727 CCATTGTAGACAGCAACACCTCCCATCATGATCATGGACACCAACCCCGATGAGGAATTC 786
QY 169 -----GLY 169
Db 787 TCCCTTAATTCATACCTGCTCAGAGCATGCTCAGSGCCCCCAGCAAGCCTCCAGTAGTGGC 846
QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrProProProProLeuSerHisAla 186
Db 847 CCTCGAACCACACAGCCAGTCAACGCTGAGGCCCCCTCTGCCACCT---CCTCATAAC 903
QY 187 HisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
Db 904 CAGACCTGTCTCCACCACCA---TCCTCTGCCAACTCCCTCAACAGAACTCCTGACC 960
QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
Db 961 AATCGCGGAGTCAAAATCCACGCCCGGCTCTCTGACCCCAATGAC-----CTGCCACC 1014
QY 223 GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuAsn 242
```

```
Db      1015  ACGCG-----GAGTCCGTTACGCTCCAGACAGCTGGGTGCTGAAC 1056
Qy      243  SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln 262
Db      1057  ACTAAGTCGCGCTGGAGCGG----- 1080
Qy      263  AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
Db      1080  ----- 1080
Qy      283  GlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
Db      1081  ---CACTTCCTCTCAAGACGTCCTCCGAAGCACACCCCTGTTCAGCAGCTTCTCCA 1137
Qy      302  GlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArgSer 321
Db      1138  GGATACCCCTTGACCTCAGGACCGTTTATACACCAACCCCGCGCTGCTGCCACGGAAT 1197
Qy      322  ThrPheAlaArgProAlaPheAsnLeuLysLysProSerIlystYrCysAsnTrpLysCys 341
Db      1198  ACATTCTCTAGGAAGGCCTTCAAGCTGAAGAAACCCCTCCAAATACTGCAGTTGGAATGC 1257
Qy      342  AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal 361
Db      1258  GCCGCCCTGTCTGCCATTGCCGTGCGCTCTCTGCGCCATTTTGTGGCTATTTCATA 1317
Qy      362  AlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu 381
Db      1318  GCAATGCATCTGCTCGGACTCAATTGGCAACTCCAGCGCGCAGATGGACACACCTTT--- 1374
Qy      382  IleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db      1375  ---AACATGGCGTAAGGACCGGCTTACCAGGAAACGATGATGTGGCAACAGTGCCA 1428
```

Search completed: August 14, 2004, 19:16:49
Job time : 864.682 secs

This Page Blank (usp)

This Page Blank (usp)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 118.339 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_1_400

Perfect score: 2143

Sequence: 1 MDVKRKPYSLRRRDAER.....EITETASSWPFVDVSLYP 400

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPRO.spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=Issued Patents NA -FMF=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=emat -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1_258@runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|----------------------|--------------------|
| 1 | 2138 | 99.8 | 1680 | US-08-891-845-3 | Sequence 3, Appli |
| 2 | 2138 | 99.8 | 1680 | US-09-514-573-3 | Sequence 3, Appli |
| 3 | 2138 | 99.8 | 3111 | US-08-891-845-1 | Sequence 1, Appli |
| 4 | 2138 | 99.8 | 3111 | US-09-514-573-1 | Sequence 1, Appli |
| 5 | 1285 | 60.0 | 2387 | US-08-891-845-11 | Sequence 11, Appli |
| 6 | 1285 | 60.0 | 2387 | US-09-514-573-11 | Sequence 11, Appli |
| 7 | 138 | 6.4 | 4446 | US-09-231-899-69 | Sequence 69, Appli |
| 8 | 131.5 | 6.1 | 1860 | US-09-051-994-1 | Sequence 1, Appli |
| 9 | 131.5 | 6.1 | 2949 | US-09-623-326-6 | Sequence 6, Appli |
| 10 | 130.5 | 6.1 | 4403765 | US-09-103-840A-2 | Sequence 2, Appli |
| 11 | 130.5 | 6.1 | 4411529 | US-09-103-840A-1 | Sequence 1, Appli |
| 12 | 130 | 6.1 | 1998 | US-09-252-991A-11748 | Sequence 11748, A |

| | | | | | | |
|------|-------|-----|-------|---|----------------------|--------------------|
| c 13 | 130 | 6.1 | 3180 | 4 | US-09-252-991A-11618 | Sequence 11618, A |
| c 14 | 128.5 | 6.0 | 8438 | 1 | US-07-945-283-1 | Sequence 1, Appli |
| 15 | 127.5 | 5.9 | 4456 | 4 | US-09-095-443-1 | Sequence 1, Appli |
| c 16 | 127 | 5.9 | 71989 | 4 | US-09-443-501A-2 | Sequence 2, Appli |
| c 17 | 126.5 | 5.9 | 43280 | 2 | US-08-804-227C-1 | Sequence 1, Appli |
| 18 | 125 | 5.8 | 2507 | 2 | US-08-471-066B-1 | Sequence 1, Appli |
| 19 | 124.5 | 5.8 | 2733 | 4 | US-03-623-326-2 | Sequence 2, Appli |
| 20 | 124.5 | 5.8 | 5121 | 4 | US-09-252-991A-15189 | Sequence 15189, A |
| 21 | 122.5 | 5.7 | 2499 | 4 | US-09-758-282B-100 | Sequence 100, App |
| 22 | 122.5 | 5.7 | 2517 | 4 | US-09-758-282B-147 | Sequence 147, App |
| 23 | 122.5 | 5.7 | 2517 | 4 | US-09-758-282B-149 | Sequence 149, App |
| 24 | 122.5 | 5.7 | 2517 | 4 | US-09-758-282B-152 | Sequence 152, App |
| 25 | 122.5 | 5.7 | 2517 | 4 | US-09-758-282B-168 | Sequence 168, App |
| 26 | 122.5 | 5.7 | 2517 | 4 | US-09-758-282B-171 | Sequence 171, App |
| 27 | 122.5 | 5.7 | 2517 | 4 | US-09-758-282B-180 | Sequence 180, App |
| 28 | 122.5 | 5.7 | 2517 | 4 | US-09-758-282B-183 | Sequence 183, App |
| 29 | 122.5 | 5.7 | 2517 | 4 | US-09-758-282B-186 | Sequence 186, App |
| 30 | 122.5 | 5.7 | 2733 | 4 | US-09-623-326-1 | Sequence 1, Appli |
| 31 | 122.5 | 5.7 | 4287 | 1 | US-08-244-189-1 | Sequence 1, Appli |
| 32 | 122.5 | 5.7 | 4287 | 1 | US-08-306-691B-53 | Sequence 53, Appli |
| 33 | 122.5 | 5.7 | 50937 | 3 | US-09-428-517-1 | Sequence 1, Appli |
| 34 | 122 | 5.7 | 4689 | 3 | US-09-105-537-34 | Sequence 34, Appli |
| 35 | 122 | 5.7 | 36778 | 3 | US-09-105-537-5 | Sequence 5, Appli |
| 36 | 122 | 5.7 | 38506 | 3 | US-09-320-878-19 | Sequence 19, Appli |
| 37 | 122 | 5.7 | 38506 | 4 | US-09-141-508-1 | Sequence 1, Appli |
| 38 | 122 | 5.7 | 38506 | 4 | US-09-657-440-19 | Sequence 19, Appli |
| 39 | 121.5 | 5.7 | 1686 | 2 | US-08-648-657-1 | Sequence 1, Appli |
| 40 | 121.5 | 5.7 | 1689 | 2 | US-08-648-657-2 | Sequence 2, Appli |
| 41 | 121.5 | 5.7 | 2517 | 4 | US-09-758-282B-174 | Sequence 174, App |
| 42 | 121.5 | 5.7 | 2727 | 4 | US-09-623-326-4 | Sequence 4, Appli |
| 43 | 121.5 | 5.7 | 2943 | 4 | US-09-379-530B-3 | Sequence 3, Appli |
| 44 | 121 | 5.6 | 4078 | 4 | US-09-016-434-1152 | Sequence 1152, Ap |
| 45 | 121 | 5.6 | 4553 | 4 | US-09-023-655-975 | Sequence 975, App |

ALIGNMENTS

RESULT 1

US-08-891-845-3
; Sequence 3, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1680 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-891-845-3

Alignment Scores:
Pred. No.: 2,49e-187 Length: 1680
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-08-891-845-3 (1-1680)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
DB 1 ATGGACGTGAGGAGGAGGAGGCTTACCGCTCGCTGACCGCGCGCGGAGCGCGC 60
QY 21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
DB 61 CGCTACACAGCTCGTCCGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaAspAlaTyrGlySerArgVal 60
DB 121 TCCAGGAGACCTTGAAGGCGCTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
DB 181 AAGGACATTGTCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
DB 241 CGGAGAGCTGGGCTGGAAGAAGTAACCGCCCTTCAAGGAGGAGGAGGAGGAGGAGG 300
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
DB 301 GGCCTCCCCACGCGCGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGAGGCTGAC 360
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
DB 361 ACGTGCTGCTCCCTCGACACCCCGTGGTCTGTGGGCGGAGGAGGAGGAGGAGGAGG 420
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
DB 421 AGCTCTGCTGCTCCAGCGGGGCAATTCATCTCACACTCACCGACACCGGAGCATGAA 480
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
DB 481 AACACTGAGACTGATCATCGGGCGGCTTCGAGAACACCGCGGCTTCGGAGCGCGCGCG 540
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
DB 541 CCGCGCTCTCGCAAGCGGACACACCCCAACAGCAGCAGCAGCGGCTCCATTAACCTCC 600
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
DB 601 AACCGGGGAACTTACGCGGAGGAGCAACCCCGAGCGCGCGCGCGCGCGCGCGCGCTC 660
QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
DB 661 TCCGAGAGCCCCCTCGCGGCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
DB 721 CTCAACAGCAACATCCCTCGGAGACCCAGAACTAGGCAAGGAGGAGGAGGAGGAGGAGG 780
QY 261 LeuGlnAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
DB 781 TTGAGGAGCAACCTCATTTAGATGGAGATTTCTCGGCGGCTTCCCGGCGGATGATGGGCTTAC 840

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
DB 841 AGTGACGGGCACTTCTCTTCAAGCTCGAGGACCTCCCGCGCTCTTCTGACCAATCA 900
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
DB 901 CCAGGGTACCCACTGACGTCAGCAGCAGTACTCTCTCCGCGCGCGCGCGCGCGCGCG 960
QY 321 SerThrPheAlaAspProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
DB 961 AGCACCTTTCGCGCGCGCGCGCGCTTTTAACTCAAGAGAGCCCTCCAGTACTGTAACTGG 1020
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
DB 1021 TCGCGAGCCCTGAGCGCCATCTGTCATCTCAGCCACTCTGTCATCTCTGTCGATACTTT 1080
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
DB 1081 GTGGCATGCACTCTTTGGCTTAACTGGCACCTGCGAGCGGATGAGGGGAGATGAT 1140
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
DB 1141 GAGATCAGGAGGACACAGCCAGGAGTGGCTGTGCCAACCGACGCTCTCCCTATACCC 1200

RESULT 2

US-09-514-573-3
Sequence 3, Application US/09514573
Patent No. 6500941
GENERAL INFORMATION:
APPLICANT: Schaefer, Gabriele M.
APPLICANT: Sliwowski, Mark
TITLE OF INVENTION: Gamma-Herregulin
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/514,573
FILING DATE: 28 FEB 2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/891845
FILING DATE: 10 JULY 1997
APPLICATION NUMBER: 60/021640
FILING DATE: 07/12/96
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P1043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1680 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-09-514-573-3

Alignment Scores: 2,49e-187 Length: 1680
Pred. No.: Alignment Scores: 2,49e-187 Length: 1680

Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-09-514-573-3 (1-1680)

Qy 1 MetAspVallyGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 1 ATGGACGTGAAGAGAGAGAAACCTTACCGCTCGCTGACCGCGCGCGCGAGCGC 60

Qy 21 ArgTyrThrSerSerAlaAspSerGluGluGluLysAlaProGlnLysSerTyrSer 40
Db 61 CGGTACACCAAGCTGCTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120

Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 121 TCCAGCGAGACCTGAAGGCTAGGACAGGACCGCGCGCTAGCTATGCGAGCGGCTC 180

Qy 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 181 AAGGACATTTGCGGAGGAGGCGGAGAGAAATCTCGCGCACAGGTGCCAACTTCACCGCTG 240

Qy 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 241 CGGAGCTGGGCTTGAAGAGAGTAACTCCCGGACCTTACCGGACCTTACCGGACAGACAT 300

Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyValadlySerAspAlaAspMetGluAlaAsp 120
Db 301 GGCTCTCCCACTGCGGCTACTCATCGGGGCTGGCTCTGATCGGACATGGAGGCTGAC 360

Qy 121 ThrValLeuSerProGluHisProValArgLeuThrGlyArgSerThrArgSerGlyArg 140
Db 361 ACGGTGTGTCCTCGACGACCGCGTGTGTGGGGCGCGAGCACACGCTCAGGGCGC 420

Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 421 AGTCTCTGCTGTCCAGCGGGGCAATTCATCTCATCTCCCGACACCGAGCATGAA 480

Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGluAsnHisAlaArgLeuArgThrProPro 180
Db 481 AACACTGAGACTGATCATCGGGGCGGCTCGAGAACCCAGCGGCTCCGAGCGCGCGC 540

Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 541 CCGCGGCTCTCGACGCGCCACACCCCAACAGCACACCGCGGCTCCATTAATCCCGCTG 600

Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 601 AACGGGGCAACTTCACGGGAGAGAACCCAGCGGCGGCGGCGGCGGCGGCGGCGGCTC 660

Qy 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 661 TCGGAGAGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 720

Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 721 CTCACAGCAACATCCCGCTGGAGACAGAGAACTTAGGCAAGCAGCGCATTCCTAGGGACA 780

Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 781 TTGAGAGCAACCTCATGTAGATGAGCATTTCTCGGCGGCTCCCGGCGGCGGCGGCTTAC 840

Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 841 AGTGACGGGCACTTCCTCTTCAAGCTGAGGAGCACCTCCCGCTCTTCTTCACACCATCA 900

Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 901 CCGGGTACCACTGACGTCCAGCACAGTGTACTCTCTCGGCGGCGGCGGCGGCGGCGG 960

Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340

Db 961 AGCACCTTCCCGCGCGGCTTTAACTCAAGAGCGCTCCAGTACTGTAACTGGAAG 1020

Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360

Db 1021 TCGCGAGCGCTGAGCGCATCTGTCATCTCAGCCACTCTGGTCATCTCTGGCATACTTT 1080

Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

Db 1081 GTGGCATGACCTGTTTGGCTTAACTGACCTGCAGCCGATGGAGGGGCGAGATGTAT 1140

Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

Db 1141 GAGATCAGGAGGACACAGCCAGTTGGCTGTGCAACCGACGCTCCCTATACCCC 1200

RESULT 3

US-08-891-845-1
: Sequence 1, Application US/08891845
: Patent No. 6096873
: GENERAL INFORMATION:
: APPLICANT: Schaefer, Gabriele M.
: APPLICANT: Sliwowski, Mark
: TITLE OF INVENTION: Gamma-Herregulin
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WinPatin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/891,845
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/021640
: FILING DATE: 07/12/96
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee, Wendy M.
: REGISTRATION NUMBER: 40,378
: REFERENCE/DOCKET NUMBER: P1043
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3111 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: US-08-891-845-1

Alignment Scores:
Pred. No.: 6,58e-187 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-08-891-845-1 (1-3111)

Qy 1 MetAspVallyGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20

Db 334 ATGGACGTGAAGAGAGAGAAACCTTACCGCTCGCTGACCGCGGCGCGAGCGC 393

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
DB 334 CGCTACACAGCTCTCGCGGACACAGCGAGGAGGCGAAAGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
DB 454 TCCAGCAGAGACCTTGAAGCCCTACACAGGAGCGCCGCTAGCCTATGGCAGCGCGTC 513
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyValAsnPheThrLeu 80
DB 514 AAGGACATTGTCCCGCAGAGGCCCGAGAAATCTCGGCACAGGTGCACAACTTCACCTG 573
QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
DB 574 CGGAGCTGGGCTGAAGAAGTAACGCCCTCACGGGACCTGTACCGGACACAT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
DB 634 GGCCTCCGCCACTGGCGCTACTCCATGGGGGCTGCTCTGATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
DB 694 ACGGTGCTGTCCCTAGACACCCCGCTGTGTGGGCGGAGCACACGGTCAGGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
DB 754 AGCTCTGCTGTCCAGCGCGGCCAATTCATCTCACACTCACCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
DB 814 AACACTGAGACTGATCATCCGGCGCGCTGCAGAACCCACCGCGCTCGGACCGCGCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAspSerIleAsnSerLeu 200
DB 874 CGCGCGCTCTCGACGCCACACCCCAACAGCACCGCGCTCCATTAACCTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
DB 934 AACCGGGCAACTTCACCGCGAGGAGCAACCCAGCGCGCGCCACCGACCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnAsnTrpLeu 240
DB 994 TCCGGAGAGCCCTCGCGCGCGCGCCAGGAGCTGCGCCACCGCCAGGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
DB 1054 CTCACAGCAATCCCTCGAGACCAAGAACTAGCAAGCAGCCATCTCTAGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
DB 1114 TTGCAGGACAACTCATTGAGATGGACATCTCGCGCGCTCCCGCCATGATGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
DB 1174 AGTGACGGGCACTTCTCTTAAGGCTTGAGGACCTCCCGCTCTCTGACCAATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
DB 1234 CCAGGATACCACTACGCTCCAGACAGTACTCTCTCCCGCCCGCCCTGCCCCG 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
DB 1294 AGCAGCTTCGCGCGCGCGCTTTAACTCAAGAAGCCCTCAAGTACTGTAACTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
DB 1354 TCGCAGGCGCTGAGCGCATCTGATCTCAGCCATCTGTCATCTCTGTCGACATCTT 1413
QY 361 VallAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
DB 1414 GTGGCATGACCTGTTGGCCATAACTGGCACCTGCAGCCGATGGAGGGGAGATGAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

DB 1474 GAGATACCGAGGACACAGCGAGCTTGGCTGTGCCAACGACGCTCTCCCTATACCC 1533
RESULT 4
US-09-514-573-1
Sequence 1, Application US/09514573
Patent No. 6500941
GENERAL INFORMATION:
APPLICANT: Schaefer, Gabriele M.
APPLICANT: Sliwkowski, Mark
TITLE OF INVENTION: Gamma-Herregulin
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/514,573
FILING DATE: 28 FEB 2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/891845
FILING DATE: 10 JULY 1997
APPLICATION NUMBER: 60/021640
FILING DATE: 07/12/96
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P1043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3111 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-09-514-573-1
Alignment Scores:
Pred. No.: 6,58e-187 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 4 Gaps: 0
US-10-029-020-14_copy_1_400 (1-400) x US-09-514-573-1 (1-3111)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
DB 334 ATGGACGTGAAGAGAGGAAGCCTTACCGCTCGCTGACCCGCGCGCGAGCCGAGCGC 393
QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
DB 394 CGCTACACAGCTCTCGCGGACACAGCGAGGAGGCGAAAGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
DB 454 TCCAGCAGAGACCTTGAAGCCCTACACAGGAGCGCCGCTAGCCTATGGCAGCGCGTC 513
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyValAsnPheThrLeu 80

Db 514 AAGACATTGTGCGGAGGCGGAGGAAATTCGCGCAGACAGTGCCAACTTCACCTG 573
Qy 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGCTGGAAGAAGTAACGCCCTCACGGGACCTGTACCGGACAGACATT 633
Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGGCTCCCCACCTGCGGCTACTCTCATGGGGCTGGCTCTGATGCGGACATGGAGGCTGAC 693
Qy 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGTGTCTCCCTGAGCACCCCGTGTCTGTGGGGCCGGAGCACACGGTCAAGGGCGC 753
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTCTGCTGTCCAGCGGGCCAAATTCATAATCTCACACTCACCGACACCGAGCATGAA 813
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATTCGGGGCGGCTGTGAGAACCCAGCGCGGCTCCGAGACGCCCGC 873
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CGCGGCTCTCGCACGCCACACCCCAACACAGCACCGCGGCTCCATTAATCTCCTG 933
Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCACGGCGAGGAGCAACCCAGCCAGCGGCGGCCACCGACCACTCGCTC 993
Qy 221 SerGlyGluProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCCCTGCGGGCGGCGCCAGAGGCTGCCACGCCAGGAGAACTGGCTG 1053
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCTCGAGACAGCAAACTAGGCAAGCAGCCATTCTTAGGAGCA 1113
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGAGGACAACTTCATTGATGATGATGATTCGCGGCTCCCGCCATGATGGGGCTTAC 1173
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCCTCTTCAAGCCCTGGAGGCACTCCCGGCTCTTCTGCACCAATCA 1233
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGTFACCACTGACGTCCAGCACAGTGTACTCTCTCGCGCCCGACCCCTGCCCGC 1293
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACTTCGCGCGCGGCTTTAACTCAAGAGGCTTCAAGTACTGTAACTGGAAG 1353
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGCGAGCCCTGAGCGGCATCGTATCTAGCCACTCTGGTCATCTCTGTCGATCTTT 1413
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGCCATGACCTGTTGGCTTAACTGGCACCTGCGAGCGGATGGAGGGGACATGTAT 1473
Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCGAGGTGGCTGTGGCAACCGGACGCTCTCCCTATACCCC 1533

RESULT 5

US-08-891-845-11
; Sequence 11, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark

; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-891-845-11

Alignment Scores:

Pred. No.: 1.66e-108 Length: 2387
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.96% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-08-891-845-11 (1-2387)

Qy 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGGCTGCGAGACCAACCGCGGCTCCGAGCGCGCGCGCTCTCGCAC 60
Qy 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCCACACCCCAACCAACAGCACCGCGGCTTCACTTAACCTCCCTGAAACCGGGGCACTTC 120
Qy 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACGCCGAGGAGCAACCCCGCGGCGCCCGAGCACCTCGCTCTCCGAGAGCCCCCT 180
Qy 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
Db 181 GCCCGCGCGCCCGAGGAGCTGCCCGCCCGAGGAGAACTGGCTGTCTCAACAGCAATC 240
Qy 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
Db 241 CCCCTGAGACCAACCAACCTAGGCAAGCAGCACTTCTTAGGAGCAATTCAGGAGCAACTC 300
Qy 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGAGCATCTCCGCGGCTCCCGCATGATGGGCTTACATGAGGAGGCACTTC 360
Qy 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305

Db 361 CTCCTTAAGCCCTGGAGGACCTCCCGCTCTTCTGCACACACATCACCGAGGTACCCACCTG 420
QY 306 ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 AGTTCAGACAGTGTACTCTCTCGCCCCGACCCCTGCCCCGAGCAGCCTTGGCCCCG 480
QY 326 ProAlaPheAsnLeuLysProSerLysTyrCysAsnTriPlysCysAlaAlaLeuSer 345
Db 481 CCGGCTTTAACTCAAGAGCCCTCAAGTACTGTAACTGGAAGTGGCAGCCCTGAGC 540
QY 346 AlalleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMethHisLeu 365
Db 541 GCATCGTCACTCAGCCACTCTGGTCACTCTGTCATCTGTCATCTTGTGGCCATGCACTG 600
QY 366 PheGlyLeuAsnTriPheHisLeuGlnProMetGluGlyGlnMetTyrCluileThrGluAsp 385
Db 601 TTGGCTTAACTGGCAGCTGCGAGCGGATGGAGGGGAGATGTATGAGATCATCGGAGGAC 660
QY 386 ThrAlaSerSerTriProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCCAGCAGTTGGCTGTGCCAACCGAGCTCTCCCTATACCCC 705

RESULT 6

US-09-514-573-11
; Sequence 11, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-09-514-573-11
Alignment Scores:
Pred. No.: 1.66e-108 Length: 2387
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 59.96% Indels: 0
DB: 4 Gaps: 0
US-10-029-020-14_COPY_1_400 (1-400) x US-09-514-573-11 (1-2387)
QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuAlaArgThrProProProLeuSerHis 185
Db 1 CATCCGGGGCGCTCGAGAACACACGCGCGCTCCGAGCGCGCGCGCTCTCGCAC 60
QY 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCCACACCCCAACACGACACACGCGCGCTCCATTAACCTCCCTGAACCGGGGCAACTTC 120
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 AGCCGAGAGAGAACCCAGCCCGCCCGCCAGCCACTCGCTCTCCGAGAGCCCTC 180
QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTriPheLeuAsnSerAsnIle 245
Db 181 GCCGGCGGCCCGAGGAGCTGCCACGCCAGGAGAACTGGCTGCTCAACAGCAATC 240
QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspLeu 265
Db 241 CCCCTGGAGACAGAAACCTAGGCAAGCAGCCATTCCTAGGAGACATTGAGGACACCTC 300
QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGGACATTTCTGGCGCTCCCGCCATGATGGGGCTTACAGTGACGGGCACTTC 360
QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 361 CTCCTTCAAGCTGGAGGACCTCCCGCTCTTCTGCACCAACATCACAGGGTACCCACTG 420
QY 306 ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 ACGTCCAGCACAGTGTACTCTCTCGCCCGCCAGCCCTGCCCGCAGCACCTTCGCCCGG 480
QY 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTriPlysCysAlaAlaLeuSer 345
Db 481 CCGGCTTTAACTCAAGAGCCCTCCAGTACTGTAACTGGAAGTGGCAGCCCTGAGC 540
QY 346 AlalleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMethHisLeu 365
Db 541 GCATCGTCACTCAGCCACTCTGGTCACTCTGTCATCTGTCGATACTTTGTGGCCATGACCTG 600
QY 366 PheGlyLeuAsnTriPheHisLeuGlnProMetGluGlyGlnMetTyrCluileThrGluAsp 385
Db 601 TTGGCTTAACTGGCAGCTGCGAGCGGATGGAGGGGAGATGTATGAGATCATCGGAGGAC 660
QY 386 ThrAlaSerSerTriProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCCAGCAGTTGGCTGTGCCAACCGAGCTCTCCCTATACCCC 705

RESULT 7

US-09-231-899-69
; Sequence 69, Application US/09231899
; Patent No. 6566583
; GENERAL INFORMATION:
; APPLICANT: Lassar, Michael
; APPLICANT: Metz, James G
; APPLICANT: Facciotti, Daniel
; TITLE OF INVENTION: SCHIZOCHYTRIUM PKS GENES
; FILE REFERENCE: CGNE.131.02US
; CURRENT APPLICATION NUMBER: US/09/231,899
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; EARLIER APPLICATION NUMBER: 09/090,793
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 4446

```
; TYPE: DNA
; ORGANISM: Schizochytrium aggregatum
US-09-231-899-69

Alignment Scores:
Pred. No.: 0.0208 Length: 4446
Score: 138.00 Matches: 106
Percent Similarity: 34.18% Conservative: 42
Best Local Similarity: 24.48% Mismatches: 163
Query Match: 6.44% Indels: 122
DB: 4 Gaps: 20

US-10-029-020-14_COPY_1_400 (1-400) x US-09-231-899-69 (1-4446)

Qy 19 GluArgGlyThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSer 38
Db 3040 GAGGTGCGGAGCTATTGCTGCAATTAAGCGCTCGCGCAAGGC-----ATC 3090
Qy 39 TyrSerSerGluThrLeuLysAlaTyrAsp-----GlnAspAlaArg 53
Db 3091 TACTCGTCGTCGACGTAAGTCTGCCCGCCGACGTCGCGCAAGCGCTCGCGGATGCCGAG 3150
Qy 54 LeuAlaTyrGlySerArgValLysAspIleValProGlnGluAla----- 68
Db 3151 TCCAGCTCGGTGCGCGGCTCGGGCATCGTTTCATGCTCGGCGTGTCTCGCGACCGT 3210
Qy 69 -----GluGluPheCysArgThrGlyAlaAsnPheThrLeuArg 81
Db 3211 CTCATCGAGAAGAGCTCCCGACGAGTTC-----GACGCGCTTTTGGCACCAG 3261
Qy 82 GluLeuGlyLeuGluGluVal-ThrProHisGlyThrLeuTyrArgThrAspIleG 101
Db 3262 GTACCGGCTCTCGAAGACCTCTCGCGCGCG-----TCGACCGC 3300
Qy 101 YLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspTh 121
Db 3301 GCCAACCTCAAGCATGGTCTCTTCAGCTCGCTCGCGGCTTCCAGCGCAAGTCCGGC 3360
Qy 121 rValLeuSer-----ProGluHisProValArgLeuTrpGly-----Ar 134
Db 3361 CAGTCTGACTACGCGCATGGCCAGCAGGCGCCCTTAACAAGATGGCGCTCGAGCTCGCCAAG 3420
Qy 134 gSerThrArgSerGlyArgSerCysLeuSer-SerArgAlaAsnSerAsnLeuThrL 154
Db 3421 GACGTCGCGTCAAGTCATGTCCTCGGTCCCTCGGACGCTGGCATGGTGCACGCGCGAG 3480
Qy 154 euThrAspThrGluHisGluAsnThrGluThrAspHisPro----- 167
Db 3481 CTCAGAAGCAGATTCCAGGAGATGGCGTGCAGATCATCCCCCGAGGCGGCGCTGAT 3540
Qy 168 -----GlyGlyLeuGluAsnHisAlaArgLeuArgThrProProProLeuSerHisA 186
Db 3541 ACCGTGGCGCG-----CATGCTGCTCGGCTCTCTCGCGGCTGAGATCTTGTGCGCAAC 3594
Qy 186 laHisThrProAsnGln-----HisHisAlaAla----- 195
Db 3595 TGGCGCACCCGCTCCAGAGGTGCGCTCGGACACCATCACCTGACCGCAAGATTTC 3654
Qy 196 --SerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAlaP 215
Db 3655 GCCAAGTCCAAACCTTCTCTCGAGGA-----CCACGT--CATCCAGGCGCGCGC 3702
Qy 215 roThrAspHis----- 218
Db 3703 GTGCTGCCCATGACGCTGGCCATTGGCTCGCTCGCGGAGACCTGCTCGGCTCTCTCC 3762
Qy 218 ----- 218
Db 3763 GGCTACTGCTCTGGCGCATTAAGCAGCGCCGAGCTCTTCAAGGGTGTCACTGTGTCAGCGC 3822
Qy 219 -----SerLeuSerGlyGluPro-----ProAlaGlyGlyAla-----GlnGluP 232
Db 3823 GACGTCAACTCGGAGGTGACCTTCAACCCGTCGACGCGCGCGCTCGGCGCGCTCAACGTC 3882
```

```
Qy 232 roAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArg----- 250
Db 3883 CAGGCCACGCTCAAGACCTTTTCCAGCGGCAAGCTGGTCCCGGCTTACCGCGCGCTCATC 3942
Qy 251 -----AsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGlu 267
Db 3943 GTGCTCTCCAAACAGGCGCGCCCGCCCAACGCGCACCATGACGCGCGCTCGCTCGAT 4002
Qy 268 MetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPhe 287
Db 4003 GCCGAT-----CCGGCGCTCCAGGCGCTCGCTTACGACGCGCAAGACCTCTTTC 4050
Qy 288 Lys---ProGly-GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuTh 306
Db 4051 CACGCGCGCGCTTCCCGGCGCATGATGCTGCTCTCGTGACCAAGACGACGCTTGTG 4110
Qy 306 rSerSerThrValTyrSerProProProArgPro-----LeuProArgSerTh 322
Db 4111 GCCAAGTGCAGCGCTGCTCCCGGCTCCGACGCGCTCGCGCGAGTTTGCACGSACT 4170
Qy 322 rPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAl 342
Db 4171 GACGCCCATGACCCCTTTCGTGAACGACCTGGCGCTTTCAGGCGCATGCTCTGTTGGTGGC 4230
Qy 342 aaLeuSerAlaIleValIleSerAlaThr 352
Db 4231 CGCAGCTCGCGCAGGCTGCGCTCCCAACT 4261

RESULT 8
US-09-051-994-1/c
; Sequence 1, Application US/09051994A
; Patent No. 6602683
; GENERAL INFORMATION:
; APPLICANT: REGENERON PHARMACEUTICALS, INC.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE EPH FAMILY LIGANDS
; FILE REFERENCE: REG-341-PCT-US
; CURRENT APPLICATION NUMBER: US/09/051,994A
; CURRENT FILING DATE: 1998-04-24
; EARLIER APPLICATION NUMBER: PCT/US96/17201
; EARLIER FILING DATE: 1996-10-25
; EARLIER APPLICATION NUMBER: 60/007,015
; EARLIER FILING DATE: 1995-10-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (202)..(1224)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1717)
; OTHER INFORMATION: n=a, c, g, or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1741)
; OTHER INFORMATION: n=a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1782)
; OTHER INFORMATION: n=a, c, g, or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1795)
; OTHER INFORMATION: n=a, c, g, or t
;
US-09-051-994-1

Alignment Scores:
Pred. No.: 0.0209 Length: 1860
Score: 131.50 Matches: 113
```

Percent Similarity: 32.35% Conservative: 41
Best Local Similarity: 23.74% Mismatches: 173
Query Match: 6.14% Indels: 150
DB: 4 Gaps: 24

US-10-029-020-14_COPY_1_400 (1-400) x US-09-051-994-1 (1-1860)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
DB 1648 CTAGACCGAGAGAGGAA-----AGATAAGAGATAGAGAGAG 1607

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
DB 1606 AAA-----ACCACTGAGAGGAGAGAGGCGGTGCCAACCAAGGTGGCGCTGT 1556

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
DB 1555 TCCCGCCCTAGCTGGGTGGGTGAGGGGAC----- 1523

QY 61 LysAsp-----IleValProGlnGluAlaGluGluPhe-----CysArgThr 74
DB 1522 AAGGATGCTGGTGTATGTCTAGAGAGAGGCTGAGCCTGTGCCCTCCCAAGGAAAAA 1463

QY 75 GlyAlaAspPheThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThr 94
DB 1462 CGATCCAGGGACACTGAGACAGAGGGGGC-----ACCCATGCCCAA 1421

QY 95 LeuTyrArgThrAspIleGly-----Leu 102

DB 1420 ACGGGAGGCGAGAACTGGGGCAGTGGGAATCCTAAGGAATCCTAAAGTGGAGACGAG 1361

QY 103 ProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu---AlaAspThr 121
DB 1360 ACAGCAGCGGAGGAGTGAAGGGCTGGGGGGCACAGGGGCCGAGATGTAGAGGTG 1301

QY 122 ValLeuSerPro-----GluHisProValArgLeuTrp----- 132
DB 1300 TCCCTCAAAACCGAATTAACCTGGAGGAGCACCCCAAGAGGGCTGGATTCAGGATAGC 1241

QY 133 ---GlyArgSerThrArgSerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsn 151
DB 1240 CAGGTGAGGAGGACCTCATACCTTGTAGTAGATGTTGGAGGGCTCTGGGGGGGGCCAT 1181

QY 152 ---LeuThrLeuThrAspThrGluHisGluAsnThrGluThrAspHisProGlyGlyLeu 170
DB 1180 CTGCACGATATACACAGGATGCCATAGTCACCACTCACCTTCTATAGTGGGGCAGA 1121

QY 171 GlnAsnHisAlaArgLeuArgThrPro-----Pro 180
DB 1120 AGGGGGGATCTGCAGCCCGCCACCCCGCAGAGACTATCCTAGCTCCCGAGGCTCAGCCT 1061

QY 181 ProProLeu---SerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSer 199
DB 1060 CCGAGGTCCATCCCACTCCACCTCCACCCCGCCAGGCGCCAGACCTCCCTCCGAGGAGC 1001

QY 200 LeuAsnArgGlyAsnPheThrPro-ArgSerAsnProSerProAlaProThrAspHisSe 219
DB 1000 CAGGACCGAGGTGGGACTCTCGAAGGCTTGGCCCGCG----- 961

QY 219 rLeuSerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTr 239
DB 960 -----TCTCCGCCAACACATGGCACCCCGCCAGCCCTGCCAGCCCGCAGCAGAGCA 911

QY 239 pLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGl 259
DB 910 GGGCAGACCCCTGCTGCTCCCGCCAGCCATGCGAGGATGCTGGG----- 868

QY 259 yThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAl 279
DB 867 -----AGGGGGCAGGGGGC 854

QY 279 aTyrSerAspGlyHisPheLeuPheLysProGlyGly-ThrSerProLeuPheCysThrT 299
DB 1785 AGGTGATTCGTGGCCCTGGACGTGGCCCTATCTCAGGGCCCTGTCTCCCTGAGGTGCCCGA 1844

853 CTTTCAGC-----ACCCCGGAGGTTGCA-----TTGCTGTGTGG 821

QY 299 hrSerProGly---TyrProLeuThrSerSerThrValTyr----- 311

DB 820 GGTACCTGGCAGGTTCCTCTCCAGGCTCCAGGCTGTGGGCTGCCCTCGGTCTCTTT 761

QY 312 ---SerProProProArg-----ProL 318

DB 760 CCATGGGCATTTCAGACACAGGTTTTCGGGGACACGCCCTCTCTCGGGACTTTGTCCCA 701

QY 318 euProArgSerThrPheAla-----ArgProAlaPheAsnLeuLysLysP 333

DB 700 CTYGAGAGAGACCTTCATGCTCTGCTTGTAGGCACACACCTCTCCTCAGGCTCTCCAGGC 641

QY 333 roSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSerAlaThrL 353

DB 640 CTTCCCGG---GTCCCATC-CGATGTGGCAATGATGTAGTAATCGTGTGGGCGAGCGAAC 595

QY 353 euValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuG 373

DB 584 TCGTGGCCCCAGAGATTAGGGCTATACTCTCTGGAACCTTGATGGT----- 541

QY 373 lnProMetGluGlyGlnMetTyrGluIleThrGluAsp----- 395

DB 540 ---GAAGCGAGATCCAGGCTCTGGCGCATCACAGTCAGAGAGGAGTTTGGGGCAGGGGT 483

QY 386 ---ThrAlaSerSerTrpProValProThrAspValSerLeu 398

DB 482 GCTCAGCGCGCGCCCTGAGACCCCTACAGGTACAGCTT 439

RESULT 9
US-09-623-326-6
; Sequence 6, Application US/09623326
; Patent No. 6607893
; GENERAL INFORMATION:
; APPLICANT: Frey et al.
; TITLE OF INVENTION: Polymerase Chimerae
; FILE REFERENCE: 4894
; CURRENT APPLICATION NUMBER: US/09/623,326
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: DE 198 10 879.6
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2949
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polynucleotide
US-09-623-326-6

Alignment Scores:
Pred. No.: 0.0432 Length: 2949
Score: 131.50 Matches: 101
Percent Similarity: 37.61% Conservative: 31
Best Local Similarity: 28.77% Mismatches: 115
Query Match: 6.14% Indels: 104
DB: 4 Gaps: 20

US-10-029-020-14_COPY_1_400 (1-400) x US-09-623-326-6 (1-2949)

QY 3 ValLysGluArgLysProTyrArg-----SerLeuThrArgArgArg 16

DB 1665 GTCAAGCAGCAGGAACTTGTAGTGGTTCTTACTTAGGAAGGCTTACGAAGAAGCA 1724

QY 17 AspAlaGluArg-----ArgTyrThrSerSerAlaAspSerGluGluGly 32

DB 1725 AGTAGTCCAAACCAAGCCCAAGTGAAGAGAGATATCAAGAGGCTCAGGAGAGCTACAC 1784

QY 33 -----LysAlaProGlnLys-SerTyrSerSerGluThrLysAlaTyrAs 49

DB 1785 AGGTGATTCGTGGCCCTGGACGTGGCCCTATCTCAGGGCCCTGTCTCCCTGAGGTGCCCGA 1844


```
Db 1333682 GCGTGTGCGGATGTTCCG-----CAGAT 1333659
QY 270 eLeuGlyAlaSerArgHisAspGly-AlaTyrSerAspGlyHisPheLeuPheLysProG 290
Db 1333658 CCGTCGCG---GAACGCCAAGATGGTTTCGTACTGATGCGCGCGTGCAGGCGCGCATGCCG 1333602
QY 290 lYgLYThrSerProLeu----- 295
Db 1333601 GCGATCGACGCAATACAGCAAAATATGTTGGTCTCGGCGTAGACCGCGGCTTCGGTCG 1333542
QY 296 --PheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProP 315
Db 1333541 GGATGTGCGCGCGCGCCCTGGGGTCCCATGCGAGCCAGAACCGCA----- 1333495
QY 315 roArgProLeuProArgSerThrPheAlaArgPro 326
Db 1333494 --CGACCTGCGGGCGCGCAACTCGCGGAAGCG 1333462
RESULT 11
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
Alignment Scores:
Pred. No.: 5,496+03 Length: 4411529
Score: 130.50 Matches: 80
Percent Similarity: 36.72% Conservative: 43
Best Local Similarity: 23.88% Mismatches: 120
Query Match: 6.09% Indels: 92
DB: 3 Gaps: 16
US-10-029-020-14_COPY_1_400 (1-400) x US-09-103-840A-1 (1-4411529)
QY 32 GlyLysAlaProGlnLysSerTyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp 51
Db 1334832 GGTCCGGTCCCAATCTGCATCGAGGTGCTCAAAATCTTTCAC----- 1334788
QY 52 AlaArgLeuAlaTyrArgSerArgValLysAspIleValProGlnGluAlaGluPhe 71
Db 1334787 -----CAGGCCAAGCCAGCGAGAT-----GACTTT 1334761
QY 72 CysArgThrGlyAlaAspPheThrLeuArgGluLeuGlyLeuGluValThrProPro 91
Db 1334760 CCGCGGCAAGTGCAACAACTTCGCGTGGGAAGCGGGATACGACGTGT-----CC 1334707
QY 92 HisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAla 111
Db 1334706 CCGTGG-----AGAACGGGA-----CGATGGTGTGGTCTCCGTGCC 1334668
QY 112 GlySer-AspAlaAspMetGluAlaAspThrValLeuSerProGluHisProVal-ArgL 131
Db 1334667 ACCAGCGAGCGGACTTTTCAC-----CTCGTCCAGCGGAATCCCAAGTCGGG 1334617
QY 131 euTrpGlyArgSerThrArgSerGlyArgSerSerCysLeuSer-----SerArgA 148
```

```
Db 1334616 CAAAGCGATGACATCGGCAACGGCGCAGCGAGCTGCTTGCCTACCTACCGTTAGCAGAT 1334557
QY 148 laAsnSerAsnLeuThrLeu-ThrAspThrGluHisGluAsnThr----- 162
Db 1334556 CGTCGAGGAACATCGCTTGAATCGGCGCGAGCCAGCAAAATCGCGTCCGCGCGCGGTG 1334497
QY 163 -----GluThrAspHisProGlyGlyLeu 170
Db 1334496 AGAGCAGGCCATACAGTTCGAGGGCAGGGAGCGGCGCGGCGCGCGGATCAGCA 1334437
QY 171 GlnAsnHisAlaArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsn 190
Db 1334436 AGTCGCCACCCAGCGCGCGGGTGCCTCCACCTGACGACGCGCGCCGCGCAAC 1334377
QY 191 Gln-HisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAs 210
Db 1334376 CAAGCGCATCAACCGCGCTAATCGCTCGCGGCGCGCGCT-----CGGCGCGACGC 1334323
QY 210 nProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaG 230
Db 1334322 CGCCGAGGACACCGCGCGGCGCACCCCGTGC-----GGCGCGCGCG 1334272
QY 230 nGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrAr 250
Db 1334271 CCAGGGTGTATGCGGCCCGCGCGCGCGCACAGGCCACCGCGCATCTT---GTCGATGCCGA 1334215
QY 250 gAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAsnLeuLeuGluMetAspIl 270
Db 1334214 CGGTGTGCGCGATGTCGCG-----CAGAT 1334191
QY 270 eLeuGlyAlaSerArgHisAspGly-AlaTyrSerAspGlyHisPheLeuPheLysProG 290
Db 1334190 CGTCGCG---GAACGCCAAGATGGTTTCGTACTGATGCGCGCTCGAGGCGCGCATGCCG 1334134
QY 290 lYgLYThrSerProLeu----- 295
Db 1334133 GCGGATCGAGCGCAATCAGACGAATATTGTTGTCGTCGCGTAGACCGCGGCTTCGGTCG 1334074
QY 296 --PheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProP 315
Db 1334073 GGATCTGCGCGCGCGCGCTGGGTGCGCATGCGAGCCAGAACCGCA----- 1334027
QY 315 roArgProLeuProArgSerThrPheAlaArgPro 326
Db 1334026 --CGACCTGCGGGCGCGCGAAGTCTCGGCGAAGCG 1333994
RESULT 12
US-09-252-991A-11748
; Sequence 11748, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11748
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11748
Alignment Scores:
Pred. No.: 0,0321 Length: 1998
Score: 130.00 Matches: 76
Percent Similarity: 33.63% Conservative: 39
```

Best Local Similarity: 22.22% Mismatches: 103
Query Match: 6.07% Indels: 124
DB: 4 Gaps: 17

US-10-029-020-14_COPY_1_400 (1-400) x US-09-252-991A-11748 (1-1998)

QY 5 GluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArgArgTyrThrSer 24
Db 913 GAGCGGCATCGCCAGCGCGT---GCTCGCGAAGCGGCGACACCGTGCACAGCTCGCG 969

QY 25 SerSerAlaaspSerGluGluGlyLysAlaProGlnLysSerTyrSerSerGluThr 44
Db 970 TCCGCTGCTGAGCATGTTCCCGCGACACCC---GTTGAGCGCGACT 1017

QY 45 LeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgValLysAspIleVal 64
Db 1018 CTACGCGCCGAGCAAGTCCATCGTTTCATCCGCGCGGCGACGCGTGTGATCCGCTA 1077

QY 65 ProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeuArgGluLeuGly 84
Db 1078 CCA-----GGCTATCC 1089

QY 85 LeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeuProGln 104
Db 1090 GTACCAAGAGTTCGGCCA-----GTACCAAGG 1116

QY 105 CysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrValLeu--- 123
Db 1117 CAAGGT-----GCAGTCGATCTCCGCGCCAGCGTCTCTATGC 1155

QY 124 ---SerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArgSerSer 142
Db 1156 CGAGCTTTCAGCATGTCGCGCGCTACCGGG-----GCTCGCGCAGGATGCGAGCA 1209

QY 143 CysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThr 162
Db 1210 GCTGTACCGCTCGGGTACCCCTCGA----- 1236

QY 163 GluThrAspHisProGlyGly-----LeuGlnAsnHis 173
Db 1237 -----CGACCAAGCGGTGACCGCTACGCGCAGCGCGTCCGCTGCAGAGCGGCAT 1287

QY 174 AlaArgLeuArgThrProPro-----ProLeu-----SerHis 185
Db 1288 GCTGCTCGAGCGCGACATCTCCAGGACACCGCGCGCTCTACGAATGGGTGCTGGAACC 1347

QY 186 AlaHisThrProAsnGlnHisHisAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 1348 GCTCTACAGCTGACCGCAAACTCTAGGAACGACCCATGGCTTTCTCGA----- 1398

QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 1399 -----CGCTCTCGCTCGCGCTGGGCGG-----CGGCT 1428

QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
Db 1429 GCCGCTGTGTCAGACCGAAGCACCAG-----ATCGGCGCTGGCGCTG 1473

QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
Db 1474 CTTGGCATGATCGCGCTACCAAGCGCACCA----- 1506

QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 1506 ----- 1506

QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 1507 ---TACCGGCTGTAGTGAACCTGCGCGCGCTCTCCGATCGCTCAAGGGCATCTCCCT 1563

QY 306 ThrSerSerThrValTyrSerProProProArgPro---LeuProArgSerThrPheAla 324
Db 1564 CAAGCACTGATCCAGACCGCCACCGCTCGGCTCGGTACCGCGCGGTGAA---GCT 1620

QY 325 ArgPro 326
Db 1621 CGACCT 1626

RESULT 13
US-09-252-991A-11618/c
; Sequence 11618, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11618
; LENGTH: 3180
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11618

Alignment Scores:
Pred. No.: 0.0668 Length: 3180
Score: 130.00 Matches: 76
Percent Similarity: 33.63% Conservative: 39
Best Local Similarity: 22.22% Mismatches: 103
Query Match: 6.07% Indels: 124
DB: 4 Gaps: 17

US-10-029-020-14_COPY_1_400 (1-400) x US-09-252-991A-11618 (1-3180)

QY 5 GluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArgArgTyrThrSer 24
Db 2314 GAGCGGCATCGCCAGCGCGT---GCTCGCGAAGCGGCGACACCGTTCACAGCTCGCG 2258

QY 25 SerSerAlaaspSerGluGluGlyLysAlaProGlnLysSerTyrSerSerGluThr 44
Db 2257 TCCGCTGCTGAGCATGCTTCCGCGCGACACCC-----GTTGACAGCGCGAACT 2210

QY 45 LeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgValLysAspIleVal 64
Db 2209 CTACGCGCGGAGCAAGTCCATCGGTTTTCATCCGCGCGGCGACGCGTGTGATCCGCTA 2150

QY 65 ProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeuArgGluLeuGly 84
Db 2149 CCA-----GGCTATCC 2138

QY 85 LeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeuProGln 104
Db 2137 GTACCAAGAGTTCGGCCA-----GTACCAAGG 2111

QY 105 CysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrValLeu--- 123
Db 2110 CAAGGT-----GCAGTCGATCTCCGCGCGCGCGCTCTCTATGC 2072

QY 124 ---SerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArgSerSer 142
Db 2071 CGAGCTTTCAGCATGTCGCGCGCTACCGGG-----GCTCGCGCAGGATGCGCGAGCA 2018

QY 143 CysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThr 162
Db 2017 GCTGTACCGGCTCGGGTAAACCTCGA----- 1991

QY 163 GluThrAspHisProGlyGly-----LeuGlnAsnHis 173
Db 1990 -----CGACCAAGCGGTGACCGCTACCGCGCGCGTCCGCTGCAGAGCGGCAT 1940

```
QY 174 AlaArgLeuArgThrProProPro-----ProLeu-----SerHis 185
Db 1939 GCTGCTGGAGCGGACATCTCCAGGACACCCGGCGCTCTACGAATGGGTGCTGGAACC 1880
QY 186 AlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 1879 GCTCTACGCTGACCGGCAAACTCTAGGAACGACCATGGCTTTCTCGA-----1829
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 1828 -----CGTCTCGCTCGGCTGGCGG-----CGCCCT 1799
QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
Db 1798 GCGGCTGGTGTCTGCAGACCAAGCCACCGA-----ATGCGGCTGGCCCTG 1754
QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
Db 1753 CTGCGGATGATCGCGGCTACCGCCACCA-----1721
QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 1721 -----1721
QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 1720 ---TACGGGCTGATGGAATGCGCGGCGCTTCTCCGTATCGCTCAAGGGCATCTCCCT 1664
QY 306 ThrSerSerThrValTyrSerProProProArgPro---LeuProArgSerThrPheAla 324
Db 1663 CAAGCAACTGATCCAGACCGCCACCGCTCGGCTCGGTATCCCGCGGTGAA---GCT 1607
QY 325 ArgPro 326
Db 1606 CGACCT 1601
```

RESULT 14

```
US-07-945-283-1/c
; Sequence 1, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The EPO and LIT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
; FILING DATE: 19920911
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ribando, Curtis P
; REGISTRATION NUMBER: 27976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 309-685-4011 ext.513
; TELEFAX: 309-685-4128
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8438 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double

;
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudorabies virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 622...6495
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1099, "g")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1267, "t")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1381, "c")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1566, "c")
; NAME/KEY: variation
; LOCATION: replace(7010, "g")
; US-07-945-283-1

Alignment Scores:
Pred. No.: 0.429 Length: 8438
Score: 128.50 Matches: 99
Percent Similarity: 36.84% Conservative: 34
Best Local Similarity: 27.42% Mismatches: 143
Query Match: 6.00% Indels: 88
DB: 1 Gaps: 17

US-10-029-020-14_COPY_1_400 (1-400) x US-07-945-283-1 (1-8438)
QY 6 ArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArgTyrThrSerSer 25
Db 5995 CGAGCCCTGCTCGACGA---GCCGCGCGCGCGAGCGGCTCGA-----TCCTCG 5948
QY 26 SerAlaAspSerGluGluGlyLys-----AlaProGlnLysSerTyrSerSerGlu 43
Db 5947 GCGGCTTGCGCGCGGAGCGCGCTCTCTCCAAACCCCAAC---TCCAGCTCCAGCTCCACC 5890
QY 44 ThrLeuLysAlaTyrAspGlnAspAlaArg-LeuAlaTyrGlySerArgValLysAspI 63
Db 5889 ACCACGTCGCGCTCGAGCCGCTCGCCCGCGCGCGCGAGAGGAGCGGACTCGGCG 5830
QY 63 eValProGlnGluAlaGluGluPheCysArg-----ThrGlyAlaAsnPheThrLe 80
Db 5829 CTCGCGGGGAGCGGCGGCGCGCTCGCAGAGACAGCCCGCGCGCGCGC-----5781
QY 80 uArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspI 100
Db 5780 -----GCCGCGGA---GAGGCGCGCTCCGCGCGCGCGC---5749
QY 100 eGlyLeuProGlnCysGlyTyrSerMetGlyAla-----GlySerAspAlaAspMetG 118
Db 5748 -----GGCTTCAGCAGCTCCAGAGCGGCGGCGGCGGAGCTCGGAGCTCTC 5705
QY 118 u---AlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrAr 137
Db 5704 CCGGCGCGCTCGCCCTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5645
QY 137 gSerGlyArgSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspTh 157
Db 5644 CTCGCGCTCTCTCTG-----TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5591
QY 157 rGluHisGluAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuAr 177
Db 5590 CGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5537
QY 177 gThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAspI 197
```



```
Db 2019 CATCCAGACAGCACCACCTCTCTCTGGATCCCAAGGTTTCCAGCCCCAAGGAT 2078
QY 252 euGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuG 272
Db 2079 TGGGCCCCAGCCCC-----AGCCCCATCTCTAGCCCCCATCTTCACAGCGT---TTGG 2129
QY 272 lyAlaSerArgHisAspGlyAlaTyr---SerAspGlyHisPhe-LeuPheLysProGly 290
Db 2130 GCCTCAGCCCCCAGCAGCGCCCTTCACACTCCAGCATCCACATCTCTTCCACCCAGGC 2189
QY 291 GlyThrSerProLeu-----PheCys----- 297
Db 2190 CCCAGGACTCTACCCCCACAATCCCCCTACCCCTATGCCCTCAGCCTGGGGTCTCTGGG 2249
QY 298 -----ThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProPro 315
Db 2250 GCAGCGCCACCCCCCTACACACCCAGCTCTACCCAGGTCCCGCTCAGACCCCTCTGCC 2309
QY 316 ArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLys 335
Db 2310 AGCCCACTCAGGGGCTCTGCCCTTCCCGAGCCCTGGGCGCCCTCAGCCTCCCATCCCC 2369
QY 336 TyrCysAsnTyrLysCys 341
Db 2370 ACTGGCATATGGTCTTGC 2387
```

Search completed: August 14, 2004, 19:39:29
Job time : 5373.34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 7074.72 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_1_400
Perfect score: 2143
Sequence: 1 MDVKERKPYSLRERRDAER.....EITETDASSWFPVDIVSLYP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_19065@runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO MMAP -IARGUEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*

- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|-----------|-------------|
| 1 | 2143 | 100.0 | 8354 | 6 | AX556500 | Sequence |
| 2 | 2138 | 99.8 | 1680 | 6 | AX270935 | Sequence |
| 3 | 2138 | 99.8 | 3111 | 6 | BD270887 | Method of |
| 4 | 2138 | 99.8 | 3111 | 6 | AR270934 | Sequence |
| 5 | 2138 | 99.8 | 3111 | 6 | BD140153 | Utilizati |
| 6 | 2138 | 99.8 | 3111 | 9 | AF009227 | Homo sapi |
| 7 | 2098 | 97.9 | 8438 | 6 | AX675551 | Sequence |
| 8 | 2082 | 97.2 | 8645 | 6 | AX600210 | Sequence |
| 9 | 2050 | 95.7 | 8585 | 10 | AB025413 | Mus muscu |
| 10 | 1971 | 92.0 | 9722 | 10 | AF059485 | Mus muscu |
| 11 | 1495.5 | 69.8 | 9264 | 5 | AB026980 | Danio rer |
| 12 | 1285 | 60.0 | 2387 | 6 | AR270936 | Sequence |
| 13 | 994 | 46.4 | 8964 | 10 | AB025412 | Mus muscu |
| 14 | 987 | 46.1 | 1476 | 9 | AF195420 | Homo sapi |
| 15 | 987 | 46.1 | 8645 | 6 | AX662355 | Sequence |
| 16 | 987 | 46.1 | 8675 | 6 | AX662353 | Sequence |
| 17 | 920 | 42.9 | 8689 | 6 | AX250067 | Sequence |
| 18 | 920 | 42.8 | 8689 | 10 | AF086607 | Rattus no |
| 19 | 917 | 42.8 | 2496 | 5 | GA245711 | Gallus ga |
| 20 | 917 | 42.8 | 2496 | 6 | AX250064 | Sequence |
| 21 | 916 | 42.7 | 8797 | 6 | AX250063 | Sequence |
| 22 | 916 | 42.7 | 8797 | 6 | AX250066 | Sequence |
| 23 | 916 | 42.7 | 8797 | 10 | AB025411 | Mus muscu |
| 24 | 897 | 41.9 | 8409 | 5 | GA279031 | Gallus ga |
| 25 | 897 | 41.9 | 8409 | 6 | AX250068 | Sequence |
| 26 | 871 | 40.6 | 8575 | 6 | AX921803 | Sequence |
| 27 | 839 | 39.2 | 9729 | 6 | AX250013 | Sequence |
| 28 | 839 | 39.2 | 9826 | 6 | AX250008 | Sequence |
| 29 | 779 | 36.4 | 8118 | 5 | GA238613 | Gallus ga |
| 30 | 736 | 34.3 | 8373 | 10 | AB025410 | Mus muscu |
| 31 | 734.5 | 34.3 | 8297 | 9 | AF100772 | Homo sapi |
| 32 | 653.5 | 30.5 | 8473 | 6 | AX662357 | Sequence |
| 33 | 653.5 | 30.5 | 8487 | 6 | AX662359 | Sequence |
| 34 | 534 | 24.9 | 1755 | 6 | AX050264 | Sequence |
| 35 | 508 | 23.7 | 1727 | 6 | AX714026 | Sequence |
| 36 | 508 | 23.7 | 1727 | 9 | AK056053 | Homo sapi |
| 37 | 506.5 | 23.6 | 7713 | 9 | HSB080325 | Homo sapi |
| 38 | 487.5 | 22.7 | 157677 | 2 | AP002470 | Homo sapi |
| 39 | 487.5 | 22.7 | 156063 | 9 | AP002515 | Homo sapi |
| 40 | 487.5 | 22.7 | 168308 | 2 | AC084775 | Homo sapi |
| c 41 | 476 | 22.2 | 87834 | 9 | AP002957 | Homo sapi |
| c 42 | 476 | 22.2 | 159946 | 2 | AC079194 | Homo sapi |
| c 43 | 458.5 | 21.4 | 93474 | 2 | AC136801 | Rattus no |
| c 44 | 458.5 | 21.4 | 173756 | 2 | AC134948 | Rattus no |
| c 45 | 458.5 | 21.4 | 245722 | 2 | AC120288 | Rattus no |

ALIGNMENTS

RESULT 1

| | | | |
|----|------|---|------|
| Db | 1 | ATGGACGTGAAGAGGAGGAAGCTTACCGCTCGCTGACCCGCGCGCGACGCCGAGCGC | 60 |
| QY | 21 | ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer | 40 |
| Db | 61 | CGCTACACAGACTCTGTCGCGACAGCAGGAGGGCAAGCCCGCAGAAATCGTACAGC | 120 |
| QY | 41 | SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal | 60 |
| Db | 121 | TCCAGCGAGACCTGAAGCCCTACGACCAAGGACGCCCGCTAGCCTATGGCAGCCGCGTC | 180 |
| QY | 61 | LysAspIleValProGlnGluIaGluPheCysArgThrGlyAlaAsnPheThrLeu | 80 |
| Db | 181 | AAGGACATTTGTGCCGACGAGGCCGAGGAATTCGCGCACAGGTGCCAACTTCAACCCTG | 240 |
| QY | 81 | ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle | 100 |
| Db | 241 | CGGAGACTGGGCTGGAGAAGTAACGCCCTCCAGGGACCCCTGACCGGACACACATT | 300 |
| QY | 101 | GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp | 120 |
| Db | 301 | GGCCTCCCCACTGGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC | 360 |
| QY | 121 | ThrValLeuSerProGluHisProValArgLeuThrGlyArgSerThrArgSerGlyArg | 140 |
| Db | 361 | ACGCTGCTGTCCCTGAGACCCCGTGGCTCTGTGGGCGCGGAGCACACGGTTCAGGGCGC | 420 |
| QY | 141 | SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu | 160 |
| Db | 421 | AGCTCTGCTGTCTCAGCGGGCCAAATTCGAATCTCACACTCACGACACCGAGCATGAA | 480 |
| QY | 161 | AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro | 180 |
| Db | 481 | AACACTGAGACTGATCATCCGGCGGCTCGAGAACACAGCGGGCTCCGAGCGCGCGC | 540 |
| QY | 181 | ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaSerIleAsnSerLeu | 200 |
| Db | 541 | CCGCGCTCTCGCAGCGCCACACCCCAACAGCACACCGCGGCTCCATTAACTCCCTG | 600 |
| QY | 201 | AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu | 220 |
| Db | 601 | AACCGGGCAACTTCAGCCGAGGAGCAACCCACGCGCGGCCCCACGACCATTCGCTC | 660 |
| QY | 221 | SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu | 240 |
| Db | 661 | TCCGGAGAGCCCTGCGCGCGCGCCAGAGAGCTGCCACGCGCCAGGAGACTGGCTG | 720 |
| QY | 241 | LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr | 260 |
| Db | 721 | CTCAACAGCAACATCCCTCTGGACACAGAAACCTAGGCAAGCAGCCATTCTTAGGACA | 780 |
| QY | 261 | LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr | 280 |
| Db | 781 | TTGAGAGACAACCTCATGTAGATGGAATTCTCGGCGCTCCCGCCATGATGGGGCTTAC | 840 |
| QY | 281 | SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer | 300 |
| Db | 841 | AGTGAGGGGCACTTCCTCTTCAAGCTTGGAGGCACCTCCCGCTCTCTCTGCACCAATCA | 900 |
| QY | 301 | ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg | 320 |
| Db | 901 | CCAGGTTACCACTGACGTCCAGCACAGTGTACTCTCTCTCGCCCGCGACCCCTGCCCCG | 960 |
| QY | 321 | SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys | 340 |
| Db | 961 | AGCACCTTCGCCCGCGCGCTTTAACTCTCAAGAAGCGCTCCCAASTACTGTAACTGGAAG | 1020 |
| QY | 341 | CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe | 360 |
| Db | 1021 | TGCGAGCCCTGAGCGGCATCTGCTACGCCACTCTGTGTCACTCTCTGTGGCATCTTT | 1080 |
| QY | 361 | ValAlaMethIstleuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr | 380 |

| | | | |
|---|--|--|-----------------|
| Db | 1081 | GTGGCCATGCACCTGTTGGCTAACTGGCACTGCAGCGATGGAGGGCAGATGAT | 1141 |
| Qy | 391 | GlulThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro | 400 |
| Db | 1141 | GAGATCAGCGAGGACACAGCCAGCTGGCTGTGCACCGACGCTCCCTATATACCC | 1200 |
| RESULT 3 | | | |
| BD270887 | | | |
| LOCUS | BD270887 | 3111 bp DNA linear | PAT 17-JUL-2003 |
| DEFINITION | Method of promoting the proliferation of inner ear hairy cells using ligand of HER2 receptor and/or HER3 receptor. | | |
| ACCESSION | BD270887 | | |
| VERSION | BD270887.1 | GI:33080655 | |
| KEYWORDS | JP 2002529425-A/6. | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| AUTHORS | 1 | (bases 1 to 3111) | |
| TITLE | Gao, W.Q. | | |
| JOURNAL | Method of promoting the proliferation of inner ear hairy cells using ligand of HER2 receptor and/or HER3 receptor | | |
| COMMENT | Patent: JP 2002529425-A 6 10-SEP-2002; GENENTECH INC | | |
| | OS | Homo sapiens (human) | |
| | PN | JP 2002529425-A/6 | |
| | PD | 10-SEP-2002 | |
| | PF | 28-OCT-1999 JP 2000580655 | |
| | PR | 07-NOV-1998 US 60/107522 | |
| | PI | WEI QIANG GAO | |
| | PC | A61K45/00,A61K38/00,A61K39/395,A61P27/16,C12N5/06/(C12N5/06,C12R1:91), | |
| | PC | A61K37/02,C12N5/00,(C12N5/00,C12R1:91) | |
| | CC | Method of promoting the proliferation of inner ear hairy cells | |
| | CC | using | |
| | CC | ligand of HER2 receptor and/or HER3 receptor | |
| | FH | Key | |
| | FT | Location/Qualifiers | |
| FEATURES | source | 1..3111 | |
| | | Location/Qualifiers | |
| | | /organism='Homo sapiens (human)' | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 1 02e-101 | Length: | 3111 |
| Score: | 2138.00 | Matches: | 399 |
| Percent Similarity: | 99.75% | Conservative: | 0 |
| Best Local Similarity: | 99.75% | Mismatches: | 1 |
| Query Match: | 99.77% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |
| US-10-029-020-14_COPY_1_400 (1-400) x BD270887 (1-3111) | | | |
| Qy | 1 | MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg | 20 |
| Db | 334 | ATGACGCTGAAGGAGGAGGAGCCTTACCGTCGCTGACCGCGCGCGGAGCGCCGAGGCGC | 393 |
| Qy | 21 | ArgTyrThrSerSerSerAlaSerGluGluGlyLysAlaProGlnLysSerTyrSer | 40 |
| Db | 394 | CGCTACACCAAGCTCGTCCGCGACAGCAGGAGGCGAAAGCCCCGCAAAATCGTACAGC | 453 |
| Qy | 41 | SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal | 60 |
| Db | 454 | TCAGGAGACCCCTGAAGGCCCTACGACCAAGGAGCGCCGCTAGCCCTATGGCAGCGCGTC | 513 |
| Qy | 61 | LysAspIleValProGlnGlnAlaGluGluPheCysArgThrGlyValAsnPheThrLeu | 80 |
| Db | 514 | AAGGACATTGTCCGCGAGGAGCGCGAGGAAATTCGCGCGACAGGTGCGCAATCACCTG | 573 |


```

QY 261 LeuGlnAspAsnLeuLeuMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACAACTCATTTAGATGACATTTCCGGCGCTCCCGCATGATGGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGAGGGGCACTTCTCTTCAAGCTTGAGGACCTCCCGCTCTTTGACACCATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGTTACCCACTGACCTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTriPlys 340
Db 1294 AGCACCTTCGCCCGCGCGCTTTAACTCAAGAGCCCTTCAAGTACTGTAACTGGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCAGCCCTGAGCGCATCTGTCATCTCAGCCACTCTGGTCTATCTCTGTCATCTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTTTGGGCTTAACTGGCACCTGCGAGCGATGGAGGGGAGATGAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGAGTTGGCTGTGGCAACCGAGCTCTCCCTATATACCC 1533

RESULT 5
BD140153 3111 bp DNA linear PAT 18-SEP-2002
LOCUS Utilization of heregulin as epithelial cell growth factor.
ACCESSION BD140153
VERSION BD140153.1 GI:23295098
KEYWORDS JP 2002509076-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 3111)
Sliwowski, M. and Kern, J.A.
Utilization of heregulin as epithelial cell growth factor
Patent: JP 2002509076-A 6 26-MAR-2002;
GENENTECH INC, UNIVERSITY OF IOWA RESEARCH FOUNDATION
OS Homo sapiens (human)
PN JP 2002509076-A/6
PD 26-MAR-2002
PF 03-FEB-1999 JP 2000530226
PR 04-FEB-1998 US 09/020598
PI MARK SLIWOWSKI, JEFFREY A KERN
PC A61K38/00, A61K35/12, A61K39/395, A61K39/395, A61P11/00, A61P43/00,
PC C07K14/485,
PC C07K16/22, C12N15/09, C12P21/08, A61K37/02, C12N15/00 CC
Utilization of heregulin as epithelial cell growth factor FH Key
Location/Qualifiers
FT source 1..3111
FT /organism="Homo sapiens (human)".
FEATURES
source
1..3111
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ALIGNMENT Scores:
Pred. No.: 1.02e-101 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x BD140153 (1-3111)

```

```

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGACAGTGAAGGAGAGAAAGCTTTACCGCTCGCTGACCCCGCGCGCGAGCGCGC 393
QY 21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCACTCGTCCGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCAGACCTTGAGGCGCTACGACCGAGAGCGCCCGCTAGCCTATGGAGCGCGCT 513
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 514 AAGGACATTTGTCGCGAGGAGGCGGAGGAAATTTGCGCGACAGGTGCCAACTTACCCCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGGGCTGGGAAGAAAGTAACGCCCGCTCACGGGACCTGTACCGGACAGACAT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCGCACTGGGGCTACTCCATGGGGGCTGGCTCTGATGCGGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTGAGCACCCGCTGCTGTGGGGCGGAGCACACGGTTCAGGGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTCTGCTGTCCAGCGGGGCAATTCCTCAATCTCACACTCACCGACACCGGACATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGGGGCTGTCAGAACCAACCGCGCGCTCCGAGCGCGCGC 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGCTCTCGCACGCGCCACACCCCCAACCCAGCACCCAGCGCGCTCCATTAACTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGGCACTTTCACCGCGAGGAGCAACCCCGCGCGCGCGCGCGCGCGCGCGCTC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTriPleu 240
Db 994 TCCGGAGAGCCCTCTGCGGGCGCGCGCGAGGCTGCCCCACCGCGAGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCGCTGGAGACAGAAACCTTAGGCAAGCAGCATTCCTTAGGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACAACTCATTTAGATGAGACATTTCTCGCGCGCTCCCGCATGATGGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGAGGGGCACTTCTCTTCAAGCTTGAGGACCTCCCGCTCTTTGACACCATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGTTACCCACTGACCTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTriPlys 340
Db 1294 AGCACCTTCGCCCGCGCGCTTTAACTCAAGAGCCCTTCAAGTACTGTAACTGGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCAGCCCTGAGCGCATCTGTCATCTCAGCCACTCTGGTCTATCTCTGTCATCTTGGCATATCTTT 1413

```

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
 Db 1414 GTGGCCATGCACCTGTTGGCTTAACCTGGACCTGCACCGGATGGAGGGGAGATGAT 1473
 QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
 Db 1474 GAGATCAAGGAGGACACAGCCAGCAGTGGCTGTGCCAACCGACGCTCCCTATACCCG 1533

RESULT 6

AF009227 3111 bp mRNA linear PRI 17-SEP-1997
 LOCUS Homo sapiens gamma-heretulin mRNA, complete cds.
 DEFINITION AF009227
 ACCESSION AF009227
 VERSION AF009227.1 GI:2406643
 KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Schaefer, G., Fitzpatrick, V.D. and Sliwowski, M.X.

TITLE Gamma-heretulin: a novel heregulin isoform that is an autocrine growth factor for the human breast cancer cell line, MDA-MB-175

JOURNAL

MEDLINE 97472144

PUBMED 9333014

REFERENCE

AUTHORS Schaefer, G., Fitzpatrick, V.D. and Sliwowski, M.X.

TITLE Direct Submission
 Submitted (18-JUN-1997) Protein Chemistry, Genentech, Inc., 460 Point San Bruno Blvd, South San Francisco, CA 94080, USA

JOURNAL

LOCATION/Qualifiers

FEATURES

Source

1..3111
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8p22-p11"
 /cell_line="MDA-MB-175"
 /cell_type="breast cancer cell"
 334..2640
 /codon_start=1
 /product="gamma-heretulin"
 /protein_id="AAC51756.1"
 /db_xref="GI:2406643"
 /translation="MDVKKPKYSLTRDAERYTSSADSECKAPKQSVSSET
 LKAYDQARLAYSRYKDIQVQPAEEFCRTGANFTREGLREVTPEHGLTRIDIGL
 PHCGYSMGAGSDADMEADTVLSEPHVRLWGRSTRSRSSCLSSRANSNLITDTEDE
 NTEDEPHGLQNHARLTTPPLSHAPTHQHAASINSNRGNFTSRNPSAPTDH
 SLGEPGAGQAPHAQENWLNINPLETRNLGKQPLGTLDQNLIEDILGSRH
 DGAYSDFHFLFKPGTSPICCTSPGVPLTSTVYPPRPLPRSTFAPAFNLKPS
 KYCNWKAALSAIVISATVILLAYFVAMHFLGNHLOPMQOMYEITEDTASWPV
 PDVSLYPSGGTGLETPDRKGTTEGKSSFPEDSFIDSGEIDVGRASOKIPRG
 FWSQVFDHPVHLKENVSLGKAALVIGYRKLGPSPHSFTQDFVELDGRRLTQBAR
 SLEGTQRQGVPPSSHETGYIYLDSDGWHIAFYNDGKSEVSVFLTAIALPRL
 KMKQSEAGSLVLRCTSSYSILRFKFNGLNKNKPQNIKIQQKPKSEL
 RINKASLDGKVMYKVIKLGNDASANTIVESNEIITGMPASTEGAVSVSEPIR
 ISVSTEGANTSSSTSTGTSTHLVCAEKEKTCVNGGBCFVWGLDLSNPRLCKP
 NEFTGRCQNVWNASYSYSTPFLSLPE"

ORIGIN

Alignment Scores:

Pred. No.: 1,02e-101 Length: 3111
 Score: 2138.00 Matches: 399
 Percent Similarity: 99.75% Conservative: 0
 Best Local Similarity: 99.75% Mismatches: 1
 Query Match: 99.77% Indels: 0
 DB: 9 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x AF009227 (1-3111)

QY 1 MetAspVallyysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20

Db 334 ATGCGCGTGAAGGAGGAGGAGGCTTACCGCTCGCTGACCCCGCGCGCGACGCGCGAGCGC 393
 QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
 Db 394 CGCTACACCACTCGTCCGCGACAGCGAGGAGGGCAAGCCCGCGAGAAATCGTACAGC 453
 QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
 Db 454 TCCAGCGAGACCTGAAGCGCTACGACAGGAGCGCCCGCTAGCTATGGCAGCGCGTC 513
 QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
 Db 514 AAGGACATTTGTCGCGAGGAGCGGAGAAATCTGCGGCACAGGTGCCAACTTCACCCCTG 573
 QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
 Db 574 CGGAGCTGGGCTGGAAGAGTAACTCCCTCAGGGACCTGTACCGGACAGACATT 633
 QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
 Db 634 GGCTCCCGCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
 QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
 Db 694 ACGGTGCTGTCCCTGAGCACCCCGTGTCTGTGGGGCGGAGACACGGTCAGGGCGC 753
 QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
 Db 754 AGTCTCTGCTGTCCAGCGCGGCGCAATTCATCTCACACTCACCGACACCGAGCATGAA 813
 QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
 Db 814 AACACTGAGACTGATCATCGGGCGGCTGCAGAACACCGCGCGGCTCCGGAGCGCGCGC 873
 QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
 Db 874 CGCGCGCTCTCGCACGCCCGCACACCCCAACACAGCACCGCGGCGCTCCATTAACTCCCTG 933
 QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
 Db 934 AACCGGGCAACTTCACGCGGAGGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCTC 993
 QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
 Db 994 TCCGGAGAGCGCCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1053
 QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
 Db 1054 CTCACAGCAACATCCCTCGGAGACAGAAACCTAGGCAAGCAGCCATTCTTAGGGACA 1113
 QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
 Db 1114 TTCCAGGACAACTCATTGAGATGGACATTCTCGCGCGCTCCCGCGCGCGCGCGCTTAC 1173
 QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
 Db 1174 AGTACGCGGCACCTCTCTTCAAGCGCTGGAGGACCTCCCGCGCTCTTCTGCACCATCA 1233
 QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
 Db 1234 CCAGGGTACCCACTGCTGACGACAGTGTACTCTCTCCGCGCGCGCGCGCGCGCGCGC 1293
 QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLeuPheSerLysTyrCysAsnTrpLys 340
 Db 1294 AGCACCTTCGCGCGCGCGCGCTTTAACTCAAGAGCGCTTCAAGTACTGTAACTGGAAG 1353
 QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
 Db 1354 TGGCGAGCGCTGAGCGCCATCGTCATCTCAGCCACTCTGTGTCACTCTCTGCTGGCATATT 1413
 QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
 Db 1414 GTGGCCATGCACCTGTTGGCCTAACTGGCACCTGCAGCGCGAGGGGCGAGATGAT 1473

```

QY      381  GluileThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
DB      1474  GAGATCAGGAGGACACAGCCAGCAGTGTGGCCTGTGCCAACCGACGCTCTCCCTATACCC 1533

RESULT 7
AX675551
LOCUS      AX675551
DEFINITION Sequence 1 from Patent WO0205704.
ACCESSION AX675551
VERSION    AX675551.1 GI:29333552
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS   Padigaru,M., Li,L., Zethusen,B.D., Casman,S.J., Shenoy,S.,
            Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M.,
            Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E., Guo,X.,
            Boldog,F.L., Grosse,W.M., Alsobrook,J.P., Gerlach,V.,
            Edingermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J.,
            Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and
            Stone,D.J.
TITLE      Proteins, polynucleotides encoding them and methods of using the
            same
JOURNAL    Patent: WO 0205704-A 1 18-JUL-2002;
            Curagen Corporation (US)
FEATURES   source
            location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      3,61e-99      Length:      8438
Score:          2098.00      Matches:    398
Percent Similarity: 98.76%      Conservative: 0
Best Local Similarity: 98.76%      Mismatches: 2
Query Match:    97.90%      Indels:    3
DB:              6      Gaps:      1

US-10-029-020-14_COPY_1_400 (1-400) x AX675551 (1-8438)

QY      1  MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
DB      4  ATGACGTGAAGAGAGAGAGAGCCCTTACCGCTCGCTGACCGCGCGCGCGAGCGC 63

QY      21  ArgTyrThrSerSerAlaAspSerGluGluLysAlaProGlnLysSerTyrSer 40
DB      64  CGCTACACCACTCGTCCGCGGACAGCAGAGGGGAAACCCCGCAGAAATCGTACAGC 123

QY      41  SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
DB      124  TCCAGCGAGACCCCTGAAGGCTACGACCAAGAGCGCCGCTAGCCTATGACGCCGCTC 183

QY      61  LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnThrLeu 80
DB      184  AAGGACATTGTGCGCGAGGAGCGCGAGGAATCTGCGCACAGGTGCCAACTTCAACCCCTG 243

QY      81  ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
DB      244  CGGAGCTGGGGCTGGAGAGTAACGCCCCCTCAGCGGACCCCTGTACCGGACAGACATT 303

QY      101  GlyLeuProGlnCysGlyTyrSerMetGlyAlaCysArgAspAlaAspMetGluAlaAsp 120
DB      304  GGCCTCCCCACTGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 363

QY      121  ThrValLeuSerProGluHisProValArgLeuThrPcLysArgSerThrArgSerGlyArg 140
DB      364  ACGGTGCTGTCCTTGGACACCCCGTGCCTGTGTGGGGCGCGAGCACACGCTCAGGGCGC 423

```

```

QY      141  SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
DB      424  AGCTCCTCCTGTCAGCGCGGCCAATTCCAATCTCACACTCACCGACACACCGAGCATGAA 483

QY      161  AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
DB      484  AACCTTGAGACTGATCATCCGGCGCGCTGCGAAGAACCGCGCGCTCCGAGCGCGCG 543

QY      181  ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
DB      544  CCGCGCTCTCGACGCCCAACACCCCAACACGACACCGCGGCTCCATTAACTCCCTG 603

QY      201  AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
DB      604  AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCGCGGCCCAACCGACCATCGCTC 663

QY      221  SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
DB      664  TCCGAGAGACCCCTGCGCGCGCGCGCCAGGAGCGCTGCCACGCCAGAGAACTGGCTG 723

QY      241  LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
DB      724  CTCACGACACATCCCTCGACACCAAACTAGCAAGCAGCCATTCTTAGGGACA 783

QY      261  LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
DB      784  TTGACGAGACAACTCATTTAGATGAGCATTCCTCGCGCTCCGCCCATGATGGGCTTAC 843

QY      281  SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
DB      844  AGTACGCGGACATTCCTCTTCAAGCCTGGAGGACCTCCCGCGCTCTTCTGCACCATCA 903

QY      301  ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
DB      904  CAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCGCGCCCGACCCCTGCCCGCG 963

QY      321  SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
DB      964  AGCACCTTCGCTGCGCGCGCTTTAACTCAAGAAAGCCCTCAAGTACTGTACTGGAAG 1023

QY      341  CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
DB      1024  TGGCGACCTTGAGCGCCATCGTCATCTCAGCCACTCTGTGTCATCTCTGCTGGCATCTT 1083

QY      361  ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet--- 379
DB      1084  GTGGCATGCACCTGTTTGGCCTAAACTGSCACCTGCGACCGATGGAGGGGCGAGTACG 1143

QY      380  ----TyrGluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLe 398
DB      1144  GATTTATGATCAGCGAGGACACAGCCAGTGTGGCCTGTGCCAACCGACGCTCTCCCT 1203

QY      398  uTyrPro 400
DB      1204  ATACCCC 1210

RESULT 8
AX600210
LOCUS      AX600210
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION    AX600210.1 GI:28400252
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS   Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S.,
            Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,
            Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K.,
            Khare,R. and Wallia,N.K.
            Proteins associated with cell growth, differentiation, and death

TITLE

```


DGAISDGHFLFKPGGTSPLFCTTSPGYPLTSTVYSPPPRLPLPRSTFRPAPNLKQPS
 KYCNWCAALGAILISATLIVLLAVFVAMHLFGLNHLHOPMEGOMQVETEDTASW
 PVPTDLSLYPSGCTGLETPDRKGAAEGKPSLFPEDSFIDSEIDVGRASQKIPP
 GTFWRQVPIHPHLKFNLSLGAALVGIYGRKGLPSHTQFDVFEULDDGRLLQTE
 ARSLGEPQGRQSPVPSSHETGFIQYLDLSGILWHLAFYNDGKSEVWSFLTJAIESVD
 NCPNCGYGDICI SGTCFLGFLGPDGCRASCPVLCSNGQYMKRCLCHSGWKGAIE
 CDVPTNQICIDVACSHGTICIMGTICINPGYKSCSEEDVMDPTCSSRGVCVRGECHC
 SVWGTCNCTPRATCLDQCSGHTFLPDTGLCNDPMSWTGDCSLEI CAADCGGHV
 CVGTGCECEDGMAACDORACHPRCAEHGTCTRDGKCSFPGWGEHCTIAHYLDVV
 KEGCEPLCNGWRCTLDLNGHVCQLGWRGTGCDTSMETGGDGKNDGDLGDCMD
 PDCCLPCLHVNPLGSPDLIIQETQAPVSOQNLNPFYDRIKFLVGRDSTHISPG
 ENPDGPHACVIRQOWTSDGTPLVGNISFINNPLFGYITSRDQSDPLVINGCISL
 ILRPERAFITQETHTPLWDRFFVMETIVMRHENETPSCDLNFPARNPVPSPSL
 TSFASCAKGPPIVPEIQALQBEIVIAAGCKMRLSVLSRTPGYKSVLRISLTHPTPE
 NLMKVLHMAVEGRFLRFAAAPDLVYFLWDTDYVKNQVFGFSEFASVGVBYES
 CPDLILWEKPRVALQGYEILDASKLGGSLDKHHLNIOGSLHKGNGFNVSQOQPV
 IGSIMGNRRRSISCPSCNGLADNKLIALVALTCSDGSLYVGDVFNFIYRFPFSNV
 TNILEMRNKDPRHSHSPAHKYKLAETDPMGSAFVLSDTNRRVFKVSTVVRKDLVKN
 EVAAGTGDQCLPFDDTRCGDGGKATEALTNPRGITVDKFLGLIYFVDGTMIRRVQNG
 IISTLLGNDLTSARPLSCDSWMEISOVRLEWPTDLA INPMDNSLYLVLDNNVLIQISE
 NHQVRI VAGRPMBHQVPGIDHFLSKVAIHATLESATALAVSHNGVLYIARTDEKKN
 RIROVTSGETLSVAGAPSGCDKNDANCDFSGDDGYAKDAKNTFSSLAVCADGEL
 YVADLGNIRIRIRKRPFLNTQNYELUSSPIDQELYFLDTSGKHLYTQSILPTGDYLY
 NFFTIGDIDITHI TDNNGNVNVRDSTGEMPLMLVDPDQVYVWYTMGTNSALRSVTQ
 GHELAMTYHGNSLLATKSNENWTTFYEDFSFGLRTNVTFTQVSSPSRSDTSV
 HVQVETSSKDDVTITNLSAGAFVLLQDOVRNSYVIGADGSLRLLLANGMEVALOT
 EPHLLAGTVNPTVKRNVTLPI DNLGLNVEWRQKEQARGQVTFGRRLVHRNRLIS
 LDPRVTRTEKIDYDHEKFTLRILLYDQAGRPSPSPSRNLGNVNTVSPSGHIAIOR
 GIMSERMEYOAGRTSIRIFADGKWSYVILEKSMVILHLSQRYIIEFEDKNDRLSSV
 TMENRVAQTLETIRSGYARNIYOPPEGNASVIOQDTEGDHLLTFIYFGKRRVYKY
 GKLSKLAETLYDTKVSFVDETAYETAMKLTNVLNQEFTCTIRYQOIGLIDRQIFRT
 EEGMNVARNFYDYNFRVTSQAVINETPLPIDLYRYDDVSGTEQFGKFGVLYDI
 NOITTAVMTHTHFHDVAGMKVEQVEI FRSLYVMVTVQYDNMGRVVKELKVGPYAN
 TTRYSYVEDADQGLQTSINDKPLWRSYDLINGNLHLLSPGNSARLPLRDLDRDRI
 RLGDVQYKMBEDGFLQRGDVFYNSAGLLIKAYNRASGWSVRYRPLRDLGRVSSKS
 SHHGLOFFYADLTNPKITVHLVHNSSEITSLYDIQGHFLFAMELSSGSEFYIACDN
 IGTPLAVFGSTGMLIKQILYAYGEIYMDTNPNFQIIIGYHGLYDPLTKLVHMRGRD
 YDLVAGRWTSPOHMLKRLSSNISVPHLYMFKNNPI SNSQIDKFCMTDVNSMLLTF
 GFLQHNVI PGVPKEDTDAMPSEYELVHTQMTQEBWDSKSLIGVQCEVQOLKAFVTL
 ERFDQVGSITISQOAPETPKPASSGSIFGKVFKPALKQDRTVIDIISVANEDGRI
 AAILNNHYLENHLITDGVDTHYFVKPGESEGLA ILGLSGGRTLENGVNVTVSOT
 NMLSGTRTRYTDQLQYRALCLNTRYGTTEDEKVRVLELARQVRAVQWAREQRL
 REGBEGLRAWTDGEKOQVLTNTRVQGYDGFVTSVEQYPELSDSANNIHPMRQSEMR
 R"

ORIGIN

Alignment Scores:

Pred. No.: 1,11e-96 Length: 8585
 Score: 2050.00 Matches: 382
 Percent Similarity: 96.52% Conservative: 6
 Best Local Similarity: 95.02% Mismatches: 12
 Query Match: 95.66% Indels: 2
 DB: 10 Gaps: 1

US-10-029-020-14_COPY_1_400 (1-400) x AB025413 (1-8585)

QY 1 MetAspValIysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
 Db 188 ATGACGCTGAGAGAGAGAGACCTATCGCTCGCTGACCCGGGCTGCGATGCTGAGCGC 247
 QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnIysSerTyrSer 40
 Db 248 CGCTACACCAAGCTCATCGGCAGACAGCGAGGGCAAGGCCCAACAGAGTCTCTACAGC 307
 QY 41 SerSerGluThrLeuIysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
 Db 308 TCACGTGAGACCTGAGGAGCATATGACCAAGATGCTCGCTAGCTACGCGACGCGCGTC 367
 QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAlaPheThrLeu 80
 Db 368 AAGGACATGATGACACAGAGGCGGAGAGTCTTCGCGCAGCAGGCACTAATTTCACTCTT 427
 QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100

Db 428 CTTGAACCTGGGACTGGGAGAGTACGCCCTCTCATGGGACTCTCTACAGGACAGACATC 487
 QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
 Db 488 GGCCTCCACACTGTGGCTATTCATGGGGGCCAGCTCTGATGACAGACTTGAAGCAGAC 547
 QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
 Db 548 ACTGTGCTATCCCTGAACCCAGTGCCTGTGGGCGCGAGCACACGGTCAAGCGCGC 607
 QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
 Db 608 AGTTCTCTCCCTGTGTCAGCGGGGCAACTCCAACTCACACTCACGACACAGAGCATGAG 667
 QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
 Db 668 AACACAGAAACCGATCATCCAGCAGCTGCAGAACCACTCGGCTCGGAGCGCTCTCT 727
 QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaSerIleAsnSerLeu 200
 Db 728 CCACACTGCCCCACGGCCATACCCCAACAGCAGCACCGCGGCTCTCATCACTCTCTG 787
 QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
 Db 788 AACAGGGGCAACTTTCACCCCAAGGAGCAACCCCAAGCCAGCACCCACAGACCTCATC 847
 QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
 Db 848 TCTGGGAGCGCCCGCAGCAGCGCGCCAGGAGCAACCCATGCGCAGGACACTTGGCTA 907
 QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
 Db 908 CTCACAGTAGTAACATCCCACTGGAAACAGAAACCTAGCAAGCAGCAGCATCTCTAGG 967
 QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuAlaSerArgHisAspGlyAlaTyr 280
 Db 968 TTGAGAGACAACTCATGTAGATGGAATCTCAGCGCCCTCCCGCCATGATGGGGCTTAC 1027
 QY 281 SerAspGlyHisPheLeuPheLysProGlyLysProGlyLysThrSerProLeuPheCysThrSer 300
 Db 1028 AGTGACGGGCACCTCTCTTCAAGCCCGGAGGAGCACCTCCCGGCTCTCTCTGCACCATCC 1087
 QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
 Db 1088 CCAGGGTACCCCTAACGCTAGCACCGTGTACTCGCCCGCACCCCGGCGCTCGCCCGC 1147
 QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
 Db 1148 AGCACCTTCTCCCGACGAGCCTTTAACTCAAGAGCCTTCCAGAGTACTGCAACTGGAAG 1207
 QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
 Db 1208 TGTGCGAGCTGAGCGCATCTCTCATCTCAGTACGCTCGCTCATCTCTGCTCGCATCTTT 1267
 QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGly-----Gln 378
 Db 1268 GTGGCCATGACACCTGTTGGCCCTAAACTGGCACCTCGCAGCGGATGGAGGGCGAGATGAC 1327
 QY 379 MetTyrGluIleThrGluAspThrAlaSerSerTyrProValProThrAspValSerLeu 398
 Db 1328 ATGTATGAGATCAGGAGGACACAGCAGCAGCTGGCTGTGTGCAACGGACGCTCTCCCTG 1387
 QY 399 TyrPro 400
 Db 1388 TATCCC 1393

RESULT 10
 AF059485
 LOCUS
 DEFINITION Mus musculus DOC4 (Doc4) mRNA, complete cds.
 AF059485
 AF059485.1 GI:3170614
 VERSION

| | | | |
|------------|---|---|-----------------------------|
| Db | 1243 | TTAGTTAAAGTCGGCTCGCCAGCGCTGGGGGTCGCAGCAACGATCATCCAGCAGCGCTG | 1302 |
| Qy | 171 | GlnAsnHisAlaArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsn | 190 |
| Db | 1303 | CAGAAACACCCCTCGGCTCCGACGCGCTCCTCCACACTGCCCCACGCCCATACCCCCAAC | 1362 |
| Qy | 191 | GlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn | 210 |
| Db | 1363 | CAGCATCACGGCGGCTCATCAACTCTTGAAACAGGGCGCAACTTACCCCAAGGAGCAAC | 1422 |
| Qy | 211 | ProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln | 230 |
| Db | 1423 | CCCAGCCAGCACCCACAGACCACTCACTCTCTGGGAGAGCCGCCAGCAGCGCCAG | 1482 |
| Qy | 231 | GluProAlaHisAlaGlnGluAsnThrPheLeuAsnSerAsnIleProLeuGluThrArg | 250 |
| Db | 1483 | GAGCAACCCATGCCAGGCAACTGGGTACTCAACAGTAAATCCCAAGTGGAAACCGA | 1542 |
| Qy | 251 | AsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIle | 270 |
| Db | 1543 | AACCTAGGCAGCAGCCATTCTTAGGCACATGGCAGCACCACTCATTTAGATGGACATT | 1602 |
| Qy | 271 | LeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysProGly | 290 |
| Db | 1603 | TTACGCCCTCCCCCGTGTAGTGGGCTTACAGTACGCGGCACCTCTCTTTTCAAGCCGGA | 1662 |
| Qy | 291 | GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrVal | 310 |
| Db | 1663 | GGCACTCCCGCTCTCTGCACCAATCCCCAGGTACCCCTTAACGCTACACCGCTG | 1722 |
| Qy | 311 | TyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeu | 330 |
| Db | 1723 | TATTGCGCCCAACCCCGGCCCTGCCCGCAGACATTCTCCCGCAGCGCTTTAACCTC | 1782 |
| Qy | 331 | LysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSer | 350 |
| Db | 1783 | AAGAGCCTTCCAACTCACTCAACTGAACTGTGACGCGTGAAGCGGCATCTCTCATCTCA | 1842 |
| Qy | 351 | AlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrp | 370 |
| Db | 1843 | GCTACGCTCGTCATCTCGTCTCGTACTTGTGGCCATGACCTGTTGGCCTAAACTGG | 1902 |
| Qy | 371 | HisLeuGlnProMetGluGly-----GlnMetTyrGluIleThrGluAspThrAlaSer | 388 |
| Db | 1903 | CACCTGAGCCGATGAGGGCGCAGATGCAGATGATGAGATCACGAGGACACAGCCAGC | 1962 |
| Qy | 389 | SerTrpProValProThrAspValSerLeuTyrPro | 400 |
| Db | 1963 | AGCTGGCCTGTGCCAACGGACGCTCCCTCGTATCCC | 1998 |
| RESULT | 11 | | |
| LOCUS | AB026980 | 9264 bp | mRNA linear VRT 05-NOV-1999 |
| DEFINITION | Danio rerio mRNA for ten-m4, complete cds. | | |
| ACCESSION | AB026980 | | |
| VERSION | AB026980.1 | GI:5307784 | |
| KEYWORDS | ten-m4. | | |
| SOURCE | Danio rerio (zebrafish) | | |
| ORGANISM | Danio rerio | | |
| REFERENCE | 1 (sites) | | |
| AUTHORS | Mieda, M., Kikuchi, Y., Hirate, Y., Aoki, M. and Okamoto, H. | | |
| TITLE | Compartmentalized expression of zebrafish ten-m3 and ten-m4, homologues of the Drosophila ten(m)/odd Oz gene, in the central nervous system | | |
| JOURNAL | Mech. Dev. 87 (1-2), 223-227 (1999) | | |
| MEDLINE | 99425191 | | |
| PUBMED | 10495292 | | |
| REFERENCE | 2 (bases 1 to 9264) | | |
| AUTHORS | Mieda, M. | | |
| TITLE | Direct Submission | | |
| Alignment | Scores: | | |
| ORIGIN | | | |
| Submitted | (06-MAY-1999) | Michihiro Mieda, Brain Science Institute, RIKEN Lab. for Developmental Gene Regulation; 2-1 Hirotsawa, Wako-shi, Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp, Tel:81-48-467-9713, Fax:81-48-467-9714) | |
| COMMENT | Sequence updated (29-Jun-1999). | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..9264 | | |
| | /organism="Danio rerio" | | |
| | /mol_type="mRNA" | | |
| | /db_xref="taxon:7955" | | |
| | /dev_stage="embryo" | | |
| gene | 1..3264 | | |
| | /gene="ten-m4" | | |
| CDS | 323..8797 | | |
| | /note="ten-m4" | | |
| | /note="similar to Drosophila melanogaster tenm/odx and Mus musculus Doc4" | | |
| | /codon_start=1 | | |
| | /product="ten-m4" | | |
| | /protein_id="BAA81893.1" | | |
| | /db_xref="GI:5307785" | | |
| | /translation="MEVKERRPYRSLTSSRRDTRRRTYSSADSEDKINPKNSYSSSET LKAFQDSRLAYGRSVKDLVHHEADFEFRQGPDFSLDMAFGDPVPPHMGAYRTEMGL PHRDYSVVSADADTETDQIMSGPEHAVLWGRNSNTKSGRSLCPSRANSNLTLTDTEH ENTENGSPVADSSASSPVDSPYPPSHAAANSOGRLLNGSAGQAQGRSEDEDFGPN SLVATGSGNVCTPAAATANEGSFQNHSLRKTPLPLFPHSHSPQHHTASIGLSLRSN YTORNPAPATDSAPNEGTSADSSAQNMMNLNSNVELETETRIAKFTLFTLEQD NFIEMDILATARDGAYTDHFLFKPGGTSPLYCTSPGYLTSVTSPPPPLRN TFSRPAFSLKPKYCHKNKCAALISLIVFLFLVIAHMLPGLANWHLPGVQVQRI YQLEBDTSGHLHPDILGLPLGNTGLEFPDRGSRDDKGLDGFPEPDEGIDMGEIVDGRVALLSGQLPGLDPPPPAQRSLVPTISHTDTCIQIMDSGIMHLAVINDGKETQV SFLTITADISIDCPNSCNFGNGVSGNCFPGFPGDCSRASCPVLCSNGQVYLKQR CMCHSGSGCECDVETNQCIDIITCSGHGTCIVGTICIMPYKGENCEVDCIDPTCSG RGVCVRGCHCFVGMWGPCSPRASCEQSGHGSFLADNTNCNDHNTWGHDCSTG LCAACDGGHGI CVAGSCRCDCRGMGTGCEORACHPRCSEHGTCKDGCCECSPGNWGEH CTIECGPLGNGRCRTLGNGWYCVCOLGMRGAGCDTSMETACSDGKDNDGDLGTLDC MDPCCLQASCHTTSLCVSGEDPDLIIQETOISLSLTSQFYQRIHLFGDSTHTVI PDVNFDFHACVIRQGVWSDGTPLVGNISFNKPAGYTITRQDGSFDLNSGVV AIGLRFEPAPITQETLMLPMWRFFVMVDTIMREVNDI PSCDLSSTREMPILVIPA PLTPAGTCRPERGIVVPIQTLQBEVRIPTGDMRLGYLSRSTSGYKLSLRLTLTHSTI PFLMKVHLWAVEGRLPKWFSAAPNLSDYFVMDKTDVYSQKVGSLSEAFVSVGFEY ESCPDLILWEKRTAVLOGYEVTASNLGWSVDKHALNIQSGILHKNGENIFISQOP VYISIMGNRRRSISCPSCNGLADKNKLLAPALACSGDSLXVGDENVYRRLFTTG NVTSVLEUSNPAHKYLIATSPVSGWLYLSDTSRKFVKSLYAVDKVAKNLELVAG TGDQCLPYDETRCGDGGKAVEATLNPRTGIVDYKGVIFFDVDTMIRIDQNGIISTL LGFNDLTARSPLCSDVMDISQVRLEWPTDLAVSPMDNSLYVDNNVVLQISENHVR I VAGRPMHCQVPGDLHFLVSKIAIHTALESANALAVSHGLYITAESEDKINRVOR STNGEISLLGAPSCGCKDANCDCYSGDGPYAKDAKLNAPSLAVSPDGLFIADL GNIRYVRNKAFLNPLNMYELISPIDDELYLFEDVNASHVFTQSLTGTGLNFTPTYS GEGDLSITIDKNKRVSLRRDSTGLPLMWLPGDQFTFTMTGNLTALKSVAQAQOETA VMTVHSGSLLATKSNEDGSGTFEYDNYGELTNVYPTGRVSYRSTDSSTVQVTE GSKNEDITVTNLASGFTYLMQDVKNYSYIIGLDSLRVLVLANGMEVSLHTPHLL SGTVNPTISKRNVLPIDINGLNVEMQREKQARGQVTVRRLRVNRNLNLMDFOR VTRTEKVYDHRKFTLRHYDHAGRPRLMAPSSRLNGVNTYS PGHLAGIQRCTMSV RMEYDNGRLTSKIFADKWSYTYLEKSWLLLYSORQIIFEPDKNDRLSSVYTMNV ARQLETRSGIYRNYRPEGNATVQDSEGLLQTHQSTGRVYIKYKGLGRVRSSTQGHH QYRLDEGFLRQNDPEFNSAGLLVKTNYKNGWTIKRYRQGLGRVRSSTQGHH LOFFYADLSFPRVTHNMHSSSETSLYLDQGLHFAMELSDEFFEDVACDNIQTL AVFSGAGLMIKQILHTAFVGLDNPFSQIVIGYGGVLEPLTKLHVHMGRRDYDLA GRWTPDHRIRKRLNSDNI VPFNLYMKNPNPLNSOETKVDNNVSLWRLVTFGLQY NVIPGYRKPVTDAPEPSYELVHTQIKTQEWNDSTKSLVGQCEVQORQLKS FVRLERQO IYASDSCGPTPLHLLTATGTSIFGKGVKVAIREGRVEADII SLANEDGRRTAAVL D KASVLDLHFTIAGLDTHYFVKGLVEGDSLIGMTVGQRTLETGTVNVTVQSVNMVGL GSRRLTIDIQMYQVGLTVYFVKGSLVGVSSVDEEKVLELARQAVATAHRLRQGBE GSRAWTDGERQQLLSGRVQGYEFTVSDQFPFELTDNINNVHFRQTEMGR" | | |

```
Pred. No.: 5,47e-68 Length: 9264
Score: 1495.50 Matches: 287
Percent Similarity: 70.34% Conservative: 45
Best Local Similarity: 60.81% Mismatches: 67
Query Match: 69.79% Indels: 73
DB: 5 Gaps: 3

US-10-029-020-14_COPY_1_400 (1-400) x AB026980 (1-9264)

QY 1 MetAspVallysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 323 ATGGAAGTGAAGAAACGACACCTTACCGCTCTCTGACCTCCAGCGGACACGAGGCGC 382
QY 21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 383 CGCTACACAGCTCATCCGCGACAGCGAGGATGCAAGATCAACCCCTAAATCTTACAGC 442
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 443 TCGAGTGAGACCTCAAGCCTTCGACAGGACTCCAGACTGGCTTATGGCAGCGCGTC 502
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 503 AAAGACCTGGTGCACCATGAGCGCGACGAGTTCAGCAGACAAGCGCGGACTTTCTCTC 562
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 563 AGAGACATCGCCCTTCGGAGTCCCGTCCGCCACATATGGGAGCGTACCGGACAGAAAG 622
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 623 GGCCCTCCCGACCGGACTACTCAGTGAGTGTGGCATCAGCGCTGACACAGAAACAGAC 682
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSer---ThrArgSerGly 139
Db 683 GGCATAATCTCCCGGACGACCGCTCGCCCTCTGGGCGCGCAGCAACCAATCCGCG 742
QY 140 ArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHis 159
Db 743 CGCAGTTCCTGCTGTTCAGCAGAGCAACTCCAACTTACGCTCACCCAGACCGAGCAT 802
QY 160 GluAsnThrGluThrAspHisPro----- 167
Db 803 GAAACACTGAGNATGGTCTCTCTGCAATGCTCATCTGCCTCTTCCCTCGTCGAC 862
QY 167 ----- 167
Db 863 TCCCTTTACCTTCCCTTCCACGACGACCAATFCAGAGCCAGGGAAGTTGTAGGTAAC 922
QY 167 ----- 167
Db 923 AGCGGGGCTCAGCGCGCGGACTCTGAGTCGGAAGATGAGTTTGGCCCAATTCATTC 982
QY 168 -----Gly 168
Db 983 TTAGTTAAACCGGCTCAGGAAACGCTGCACTCCGCGCGCTACTGTAATGAGGCG 1042
QY 169 GlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHisAlaHisThr 188
Db 1043 TCGTTCCAGAAATCACTACGTTTGCAGAACGCCACCAATTAACCTGTTTCACTCGCACTCC 1102
QY 189 ProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArg 208
Db 1103 CCAGTGCAGCACACACCGGCTCATCGGTTCCCTGAGCGCGCAGCACTACACCCAGCGC 1162
QY 209 SerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProAlaGlyGly 228
Db 1163 AGTAACCGGAGCGCGGACCCACCGATAGTTCAGTCCCAATAGGTTCCGACA---AGT 1219
QY 229 AlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGlu 248
Db 1220 GCCCAGGACTCCAGCGCGCGGACCACTGGCTTCTCAACAGCAACGTTCTCTAGAG 1279

249 ThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMet 268
Db 1280 ACCAGAAACATAGCAAGACAGACATCTCTAGAGACTTTGCAGGACAACTTTATAGAGATG 1339
QY 269 AspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLys 288
Db 1340 GACATACTGGCGACAGCTCGCGCGACCGGCTTACACTGATGGACACTTTCTTTTCAAG 1399
QY 289 ProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 1400 CCTGGCGGAGCTCGCCCTTTGTACGACCACTTCACTCGTGTACCTTGTACGTCGACG 1459
QY 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 1460 ACGGTGTACTCTCTCCCGCGCGGCTGCGCAAGAACACCTTCTCAGCGGCTGCCTTT 1519
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1520 AGCCTAAAGAAACCCCTATAAGCACTGCAACTGGAATGTGCGCCCTCAGTGCATCTC 1579
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1580 ATCTCGCTCACCTCGTGTCTCTGTGGCATACTTTATTTGCCATGCACCTCTTTGACTC 1639
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSer 388
Db 1640 AACTGGCATTTGACGCGGTGAGAGGAGGAGATATATCATGCTGACGAGGACAAACCACT 1699
QY 389 SerTrpProValProThrAspValSerLeuTyrPro 400
Db 1700 GGCCTACATCTGCCACAGATCTGGGCTGCCACCT 1735

RESULT 12
LOCUS AR270936 2387 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 11 from patent US 6500941.
ACCESSION AR270936
VERSION AR270936.1 GI:29702178
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE 1 (bases 1 to 2387)
AUTHORS Schaefer,G.M. and Sliwkowski,M.
TITLE Gamma-heredulin
JOURNAL Patent: US 6500941-A 11 31-DEC-2002;
FEATURES
source Location/Qualifiers
1..2387
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: Length: 2387
Pred. No.: 9,1e-58 Matches: 235
Score: 1285.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 59.96% Gaps: 0
DB: 6

US-10-029-020-14_COPY_1_400 (1-400) x AR270936 (1-2387)

QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGGCTCGCAGAACACACCGCGGCTCCGACGCGCGCGCGCGCTCTCGCAC 60
QY 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCCAGACCCCGGACCGACCGGCGCTCCCTTAACCTCCCTGAAACCGGCGCACTTC 120
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACSCCGGAGGAGCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 180
```

```

QY 226 AlaGlyGlyAlaGlnGlnProAlaHisAlaGlnGlnAsnTrpLeuLeuAsnSerAsnIle 245
Db 181 GCCCGCGCGCCAGGAGCCTGCCCGCCAGGAGAACTGGCTGCTCAACAGCAACATC 240
QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
Db 241 CCCCTGAGACCAAGAACTTAGGCAAGCAGCATCTCTAGGGAATTCAGAGCAACCTC 300
QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGACATTCCTCGCGCCTCCGCCCATGATGGGCTTACAGTGACGGGCACTTC 360
QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 361 CTCTTCAAGCCTTGAGGACACCTCCCGCTCTCTTCTGCACCATCACCAGGGTACCCACTG 420
QY 306 ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 ACGTCCAGCAGTGTACTCTCTCCGCCCGCAGCCCTTCCGCCCGCAGACCTTCGCCCGG 480
QY 326 ProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLysCysAlaLeuSer 345
Db 481 CCGGCCCTTAACCTCAAGAGCCCTCCCAAGTACTGTAACCTGGAAGTGGCAGGCCCTGAGC 540
QY 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisIleu 365
Db 541 GCCATCGTCATCTCAGGCATCTCGTCATCTCGTCGCATACCTTGTGGCCATCAGCCTG 600
QY 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
Db 601 TTGGCCTTAACCTGGCACCCTGCAGCCATGAGGGGAGATGATGAGATCAGCGAGGAC 660
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCAGCAGTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 705

```

RESULT 13

```

AB025412 8964 bp mRNA linear ROD 08-MAY-1999
LOCUS Mus musculus mRNA for Ten-m3, complete cds.
DEFINITION AB025412
ACCESSION AB025412
VERSION AB025412.1 GI:4760779
KEYWORDS Ten-m3
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Cohashi T., Zhou X., Feng X., Richter, B., Moergelin, M., Perez, M.T.,
Su, W., Chiquet-Ehrismann, R., Rauch, U. and Faessler, R.
Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
J. Cell Biol. (1999) In press
2 (bases 1 to 8964)
Direct Submission
Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
(E-mail: oohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128,
Fax: +81-86-222-7768)
Location/Qualifiers
1. 8964
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balt/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1. 8964
/genes="ten-m3"
50. 8197

```

FEATURES

```

source
1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
50 ATGGATGTGAAGAAACGACGACCTTACTGCTCTCTGACCAAGAGACAGCGGAAAGGAA 109
20 ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
110 AGCGCTATACAAATTCGTCGCGGACATGAGAGATGAGGCTCCCGACGAGAGACTCC 169

```

```

/gene="ten-m3"
/codon_start=1
/product="Ten-m3"
/protein_id="BAA77398.1"
/db_xref="GI:4760780"
/translation="MDVKRRPYCSLTKSRREKRRVYTNSSADNEECRVPTQKSYSS
ETUKAFDHYSLRLYGNRKVDLVRHDEAYTRQGNFTLRQLGVCESATRGVAFCAE
MGLPHRGYSISAGSDADTENEAVMSPEHMLWGRGVKPSGRSSCLSSRSNLNRNLTDT
EHNRSDESSEQSNPNPQPTLOPLPPSHKHQHPAQHHSITSLSNRNYSNTRNOSPAP
PAALPAELOTPPESVOLQDSNVLPSNHLFVKTGTTTFLFSTATPGYTMAAG
SVYSPPTPLPRNTLSAFAFKKSKSCSWRCTALCAVGSVLLAILLSYFIAMHLF
GLNWHLOQTENDTFENGKNSDTPVNTVSLPSGDNKLGFTHENNITDELGDLGR
RAIQEVPFGLFWSQFLDQFLKFNLSIQKDALIGVYGRKGLPPSTQYDFVELLD
GSLRIAREQNLNRESRAGQARSVLSHEAGFTQYLDGIMHIAFVLDGKNGKQVSEFN
TIVESVVECPNCHNGECVSGTCHCFPGFLGDCSRAACPVLCVNGDQVSKGKCLC
FSWGKTECPVPTTQCIDPOCGRGICIMGSCACNCGVKGENCEADCLDPCGSHGV
CIGHECHCNPGWGSCEILLTWCADOCGSGHTYLOBSGCTCDPMTWGTGDCNEICS
VDCSGHGVCMGSCRCRCEGWTGPACNORACHPCRAEGTKCDKGKCECQGWNGEHTI
AHYLDKLVKEGCPGLNSNGRCLTDQNGHVCPCQWRGAGCDVAMETLCTDSKDNHG
DLIDMDPCCLQSSCQNPYCRGLPDQDIISQSLQTPSQAAKSYFDRISFLIGS
DSTHVLPGESPFNKLASVIRGVLTAADGTPGLIGNVSYFLHSEYGYTITROQMFOL
VANGASGLTLVFFERSPLTOYHTVMI PNWVYVMDTLVMKEENDI PSCDLSGKVRFS
PIIVSPLSTFERSSPDSPIIPETOVLHEETIPGTDLKLSYLSAQAQVSKVLKT
MTQAVIPNLMKHLMAVAVVGRLEFQKWFPSPNLAYTIFINDTDAYNQKYVGLSEAVV
SVGTESCDLTLWEKRTAVLOQYELDASNGMGTLDKHHVLDVQNGILYKNGENYV
FISQOPEVSSIMNGRRRSISCPSCNQADGNKLLAPVALACIDGSLYVGDFNYVR
RPFSGNVTSLVLENRKDFRHSNPARYLATDPVTGDLVSDTNTTTRRYRPSKLTG
AKDLTKNAEVVAGTEOCLPDEARCGDGKAVEATLMSPKGMALDKNGLTYFDVTGM
IRVQDNGIISTLLGNDLTSARPLCTDCTSMHISOVRLEWPTDLAINPMDSILVLDN
NVUOLITNRPQVIAAGRPMHCQVPGVEYVPGKHAVQTTLTSAITAVISGVLYITE
TDEKKINRQVTTDGEISLVAGIPSECDCKNDANDCYOSGDGYAKADKLAPSSLA
ASPDGLTYIADLNIRIRAVSKNKLPLNSMNFVYASPTDOELYIFDINTHOYTVSL
VTGDIYNSYNDNDVTAVDNSGNTLRIRRDENRMPVRVSPDNQVILWTGTNGC
LKMTAQLLEVLFTVHSGLLATKSDTGTTFDYDSEGLTNTVTFPGVVTNLH
GMDMKAITVDIESSRDEVSITNSLSSIDSFTVMDQOLRNSYVOIGDGLRFLYAS
GLDSHYOTEPHLVACTANTPAKNTMLPGENQNLVEMFRFEQAGCKVNVGRKLR
VNGRNLSDFDRTTKETIKYIDHRRKFLRLAIADTSGHTLMLPSSKLMVAVTYSST
GQIASIQGTTSEKVDYDSQRIYRVFADGKTWSTYLTLEKSMVLLHSQRQYIFEDY
MMDRLSAITMPSVARHTMTQIRSIGYRNTINPPESNASITDYNREGLLQTLAFLGT
SERLVFKYRQTRLSEILLYDSTRVSFTYTAGVLTKNVLSQSDGFTCYIRQIGPLI
DROIIRFSDGMVNARFDYSYDNSFRVTSMQVNETPLPDLQPDIDLSKVQFOFK
FGVIYDINQLISTAVMTYTKHFDAGRIKELQYEIFRSLMYMTIYDQNMGRVTKRE
IKIGPANTIKYAYEYDVQQLQTVILNKKIMRYNDLNLHLLHLLHLLHLLHLLHLLH
YDLRDRITRIGVDVQRLDEDFLQRQETFEYSSKGLLTRYYSKSGWTVIYRYDGL
GREVSKTSILGHLQFFYADLTPTIRTHVNHSSSEITLKYLDLQGLFAMETISGD
EFYASDNTGTPLAFTSSNGLMKQYTAIGEIFYFDSNVDFQLVIGFHLGXDPLTK
LHIFGRDVIDLAGRWTTDPIETWKRIKDPAPNLMYRNPNPASKIHVDKDYITDV
NSLWTFGPHLNAIPGPVPKPDLTSPVELKVSQOWEDVPPIFGVQQQVQAQKAF
LSLGMKAEVQVSRKAGAQSWLFAFVKSLLKGKVMVAVSQGRVQTNUNIANEDCI
KVAALNNAFYLENLHFTLEKDTHTYFKITTPESDLGTLRLTSGRKALENGINVTVS
QSTVTVNGTRTRFADVMQFALALHVRVGMTLDEEKRIEQAQRLAENAREEQ
RVSDGEGARLWTEGEKRLSAGKVQGVYVLSVEQPELADANNITQFLRQSEI
GKR"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 4,35e-42 Length: 8964
Score: 994.00 Matches: 212
Percent Similarity: 63.13% Conservative: 50
Best Local Similarity: 51.08% Mismatches: 89
Query Match: 46.38% Indels: 64
DB: 10 Gaps: 12

```

US-10-029-020-14_COPY_1_400 (1-400) x AB025412 (1-8964)

```
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 170 TATAGTTCCAGTGAACCTTGAAGCTTCGATCATGATTAATTCACGGCTGCTTTATGGA 229
QY 58 SerArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaVal 77
Db 230 AACAGAGTAAGGATTTGGTCCACAGAGAAAGCCGACGATGATCTAGACAGAGAGAAAT 289
QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97
Db 290 TTTACCTTAAGCGATGAGGTGTGTGAATCCGCAACTCGAAGAGGAGTGGCATCTGTG 349
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 350 GCGGAATGGGCTCCCTCACAGAGGTACTCCATCAGTGCAGGTGATGATCGGATAGC 409
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137
Db 410 GAAACAGAGAGCAGTATGTCCTGAGCATGCCATGAGACTTTGGGGCAGGGGGGTCAA 469
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 470 TCGGGCCGAGTTCCTGCTGTCAGCGGTCCAACTCCGCCCTCACCTGACAGACAG 529
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 530 GAGCAGAGAACAGTTCGAGCAGTGCAGAGCAGCAACTTCAACCAACCCAGGCAACCC 589
QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisAla 194
Db 590 ACCCTGCGACCTTTCGCGCA-----TCCCAAGCAGCAGCACCAGCGGATCAC----- 640
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 641 CGCTCATCACTTCCCTCAATAGAACTCCCTGACCAATAGAAAGAACAGAGTCCGGCC 700
QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 701 CGG-----CGGGTGTCTTTCGCGCGAGCTGCAAAACCA 736
QY 231 ---GluProAlaHisAlaGlnGluAsnThrLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 737 CCGGAGTCCGTCAGCTGCAGACAGCTGGCTTGGCAGTAATGTACACTGGAAGC 796
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 797 AGS----- 799
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPhelysPro 289
Db 800 -----CAATTCCTATTCAAAACA 817
QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 818 GGGACAGGACGACGACCTGTCAGTACGCAACCCCGGATACCAATGATGATCGG 877
QY 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 878 TCTGTTTATCTCCGCTACCCGCGCACTTCTAGAAACACCCCTATCAAGAGTGTCTTT 937
QY 329 AsnLeuLysLysProSerLysTyrCysAsnThrLysCysAlaAlaLeuSerAlaIleVal 348
Db 938 AAATTCAGAAAGTCTTCAAAGTACTGACGTGGAGGTGCACCGCACTGTGTGTGAGGG 997
QY 349 IleSerAlaThrLeuValIleLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 998 GTCTCAGTCTCTCGGCACTTCTCCTCTCTATTTATAGCAATGCACTATTGCGCTC 1057
QY 369 AsnThrHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1058 AACTGCACTTACAGCAGACGGAAATGACACATTCGAGATGAAAGATGAATCTGTGAC 1117
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
```

```
Db 1118 ACC-----GTGCCAACAACACTGTATGTTACCT 1147
RESULT 14
AF195420
LOCUS AF195420 1476 bp mRNA linear PRI 27-JAN-2000
DEFINITION Homo sapiens ODZ3 (ODZ3) mRNA, partial cds.
ACCESSION AF195420
VERSION AF195420.1 GI:6760372
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1476)
AUTHORS Ben-Zur, T., Feige, E., Motro, B. and Wides, R.
TITLE The mammalian Odz gene family: Homologs of a Drosophila pair rule
gene with expression implying distinct yet overlapping
developmental roles
JOURNAL Dev. Biol. 117, 107-120 (2000)
REFERENCE
2 (bases 1 to 1476)
AUTHORS Ben-Zur, T., Motro, B. and Wides, R.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-1999) Faculty of Life Sciences, Bar-Ilan
University, Ramat-Gan, Israel
FEATURES
Location/Qualifiers
source
1..1476
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>1476
/gene="ODZ3"
/note="Ten-m3"
151..>1476
/gene="ODZ3"
/note="similar to the Drosophila melanogaster pair rule
odd Oz/Ten-m gene sequence"
/codon_start=1
/product="ODZ3"
/protein_id="AAF28318.1"
/db_xref="GI:6760373"
/translation="MDVKRRPYCSLTKSRRKERYTNSADNEECRVPTOKSYSSS
ETLKAFDSSSLYGNRVKDLVHREADFTRQGNFTRQLGVCPEATRRGLAFCAE
MGLPHRYSISAGSDADTENEAMVPEHMLWGRVSKSSRSSLSRNSALTUTDT
EHENKSDNEQPSNQCSLTQLPPSHKQSAHQHPSITSLNRSITNRRNQSPAP
PAALPAELQTTPESTVQLQDSWLVGNVPLESRHFLTKTGTPTPLFSTATPGYTWASG
SVYSPPTPLPNTLSRAFKPKSKSKYCKWCTALCAVGSVLLAILLSYFIAMHLF
GLNWLOQTENDTPENGKVNSDTMTNTVSLPSGNGKLGFGTQENNTIDSGELDIGR
RAIQEIPFPFRSOLFIDQPFQFLKFNISLQKDALIGVGRKKLPESHQ"
```

Alignment Scores:

| | | | |
|------------------------|----------|----------------|------|
| Align. No.: | 1,34e-42 | Length: | 1476 |
| Score: | 997.00 | Matches: | 212 |
| Percent Similarity: | 62.8% | Conservatives: | 49 |
| Best Local Similarity: | 51.0% | Mismatches: | 90 |
| Query Match: | 46.0% | Indels: | 64 |
| DB: | 9 | Gaps: | 12 |

US-10-029-020-14_COPY_1_400 (1-400) x AF195420 (1-1476)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19

Db 151 ATGGATGTGAAGAGCGCGCTTACTGCTCCCTGACCAAGAGCAGCAGAGAGAA 210

QY 20 ArgArgTyrThrSerSerAlaAspSerGluGlyLysAlaPro---GlnLysSer 38

Db 211 CGCGGTACACAAATCTCCGACACATGAGAGTGCCTGACCAAGAGTCC 270

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57

Db 271 TACAGTTCAGCGAGACATTTGAAGCTTTTGATCATGATTCCTCGCGGTGCTTTACGGC 330

```
Qy 58 SerArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsn 77
Db 331 AACAGAGTGAAGATTGGTTTCACAGAGACAGACGAGGTTCATAGCAAGGACAGAAAT 390
Qy 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTTACCTTAAGCGAGTTAGGAGTTGTGAACAGCACTCGAAGAGGAGTGGCATTTGT 450
Qy 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGAATGGGGCTCCCTCACAGAGGTTACTCTATCAGTCAGGGTCAGATGCTGATACT 510
Qy 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137
Db 511 GAAATAGAGCAGTGATGCTCCAGAGAGTCCAGAGCTGAGACATTTGGGGCAGGGGTCAA 570
Qy 138 SerGlyArgSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCGCAGCTCTGCTCTCAAGTCGGTCCAACTCAGCCCTCAGCCCTGACAGATACG 630
Qy 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAACAAAGTCGACAGTGAGAAATGAGCAACCTGCAAGCAATCAAGGCGAGTCT 690
Qy 175 ArgLeuArgThrProProLeuSerSerHisAlaHisThrProAsnGlnHisAla 194
Db 691 ACCCTGCAGCCCTTGGCGCT-----TCCATAGCAGCACTCTGCACACATCAT--- 741
Qy 195 AlaSerLeuAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCATCACTTCTCTCAACAGAAACTCCCTGACCAATAGAGGAACACAGAGTCGGCC 801
Qy 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CCGCGCTCTTTGGCCGCGAGCTGCAAAACACACA 837
Qy 231 ---GluProAlaHisAlaGlnGluAsnTripleLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCCGAGTCCGTCAGTCGACGAGCAGTGGTCTTGGCAGTATGATACCACTGGAAAGC 897
Qy 250 ArgAsnLeuGlyLysGlnProPheLeuLeuGlyThrLeuGlnAspAsnIleGluMetAsp 269
Db 898 AGG----- 900
Qy 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTTCTTATTCAAAACA 918
Qy 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGGTACAGCCCACTGTTTCAGTACTGCAACCCAGAGATACAAATGGCATCTGGC 978
Qy 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTACACCACCTACTCGGCCACTACCTAGAAACACCCCTATCAAGAGTGCTTT 1038
Qy 329 AsnLeuLysLysProSerLysTyrCysAsnTriPysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAAGAAAGTCTTCAAGTACTGAGCTGGAAATGCATGCACTGTGTGCGGTAGGG 1098
Qy 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCGGCAATCTCTCTTATTTATAGCAATGATCATCTCTTGGGCTC 1158
Qy 369 AsnTriPHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGCTACAGCAGACTGAAATGACACATTTGAGAATGAAAAGTGAATTTCTGAT 1218
Qy 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1219 ACC-----ATGCCAAACAAACACTGTGTCTTACCT 1248
```

RESULT 15

```
AX662355
LOCUS AX662355 8645 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 37 from Patent WO02062999.
ACCESSION AX662355
VERSION AX662355.1 GI:29163217
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS
Szytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
Malyan, U., Shinkets, R.A., Tchernev, V.I., Spaderna, S.K.,
Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
Millett, I. and Macdougall, J.R.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 02062999-A 37 15-AUG-2002;
Curagen Corporation (US)
FEATURES
source
1..8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 9.6e-42 Length: 8645
Score: 987.00 Matches: 212
Percent Similarity: 62.89% Conservative: 49
Best Local Similarity: 51.08% Mismatches: 90
Query Match: 46.06% Indels: 64
DB: 6 Gaps: 12
US-10-029-020-14_COPY_1_400 (1-400) x AX662355 (1-8645)
```

```
Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGATGTGAAAGAAACGACGAGCCCTTACTGCTCCCTGACCAAGAGCAGCAGAGAGAA 210
Qy 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGGCGCTACACAAATTCCTCCGACACATGAGAGTGCCGGGTACCCACAGAGTCC 270
Qy 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTTCACGAGACATTTGAAAGCTTTTGTATCATGATTCCTCGCGCTGCTTTACGGC 330
Qy 58 SerArgValLysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsn 77
Db 331 AACAGAGTGAAGATTGGTTTCACAGAGACAGCAGGATTCATAGCAAGGACAGAAAT 390
Qy 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTTACCTTAAGCGAGTTAGGAGTTGTGAACAGCACTCGAAGAGGACTGGCATTTGT 450
Qy 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGAATGGGGCTCCCTCACAGAGGTTACTCTATCAGTCAGGGTCAGATGCTGATACT 510
Qy 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137
Db 511 GAAATAGAGCAGTGATGCTCCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 570
Qy 138 SerGlyArgSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCGCAGCTCTCTGCTTCAAGTCGGTCCAACTCAGCCCTCAGCCCTGACAGATACG 630
Qy 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAACAAAGTCCGACAGTGAGAAATGAGCAACCTGCAAGCAATCAAGGCGAGTCT 690
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 12:05:42 ; Search time 5140 Seconds
(without alignments)
3372.999 Million cell updates/sec

Title: SEQ14-X-AT-28-64-76
Perfect score: 2132
Sequence: 1 MDVKRPYSLTRRDAER.....EITEDTASSWPVPTDVSILP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/MITRA020/runat_06082004_114100_206/app_query.fasta.1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPGL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MITRA020 @CEN_1_1_3731_@runat_06082004_114100_206 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rodi.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htg.hum.*
40: em.htg.mus.*
41: em.htg.other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 2127 | 99.8 | 8354 | 6 | AX556500 Sequence |
| 2 | 2122 | 99.5 | 1680 | 6 | AR270935 Sequence |
| 3 | 2122 | 99.5 | 3111 | 6 | BD270887 Method of |
| 4 | 2122 | 99.5 | 3111 | 6 | AR270934 Sequence |
| 5 | 2122 | 99.5 | 3111 | 6 | BD140153 Sequence |
| 6 | 2122 | 99.5 | 3111 | 9 | AF009227 Homo sapi |
| 7 | 2082 | 97.7 | 8438 | 6 | AX675551 Sequence |
| 8 | 2066 | 96.9 | 8645 | 6 | AX600210 Sequence |
| 9 | 2038 | 95.6 | 8585 | 10 | AB025413 Mus muscu |
| 10 | 1959 | 91.9 | 9722 | 10 | AF059485 Mus muscu |
| 11 | 1482.5 | 69.5 | 9264 | 5 | AB025413 Mus muscu |
| 12 | 1285 | 60.3 | 2387 | 6 | AR270936 Sequence |
| 13 | 982 | 46.1 | 8964 | 10 | AB025412 |
| 14 | 975 | 45.7 | 1476 | 9 | AF195420 Homo sapi |
| 15 | 975 | 45.7 | 8645 | 6 | AX662355 Sequence |
| 16 | 975 | 45.7 | 8675 | 6 | AX662353 Sequence |
| 17 | 905 | 42.4 | 2496 | 5 | GGA245711 Gallus ga |
| 18 | 905 | 42.4 | 2496 | 6 | AX250064 Sequence |
| 19 | 904 | 42.4 | 8689 | 6 | AX250067 Sequence |
| 20 | 904 | 42.4 | 8689 | 10 | AF086607 Rattus no |
| 21 | 904 | 42.4 | 8797 | 6 | AX250063 Sequence |
| 22 | 904 | 42.4 | 8797 | 6 | AX250066 Sequence |
| 23 | 904 | 42.4 | 8797 | 10 | AB025411 Mus muscu |
| 24 | 885 | 41.5 | 8409 | 5 | GGA279031 Gallus ga |
| 25 | 885 | 41.5 | 8409 | 6 | AX250068 Sequence |
| 26 | 859 | 40.3 | 8575 | 6 | AX921803 Sequence |
| 27 | 828 | 38.8 | 9729 | 6 | AX250013 Sequence |
| 28 | 828 | 38.8 | 9826 | 6 | AX250008 Sequence |
| 29 | 777 | 36.4 | 8118 | 5 | GGA238613 Gallus ga |
| 30 | 734 | 34.4 | 8373 | 10 | AB025410 Mus muscu |
| 31 | 733.5 | 34.4 | 8297 | 9 | AF100772 Homo sapi |
| 32 | 641.5 | 30.1 | 8473 | 6 | AX662357 Sequence |
| 33 | 641.5 | 30.1 | 8487 | 6 | AX662359 Sequence |
| 34 | 522 | 24.5 | 1755 | 6 | AX050264 Sequence |
| 35 | 508 | 23.8 | 1727 | 6 | AX714026 Sequence |
| 36 | 508 | 23.8 | 1727 | 9 | AK056053 Homo sapi |
| 37 | 506.5 | 23.8 | 7713 | 9 | HSM808325 Homo sapi |
| 38 | 487.5 | 22.9 | 157677 | 2 | AP002470 Homo sapi |
| 39 | 487.5 | 22.9 | 166063 | 9 | AP002515 Homo sapi |
| 40 | 487.5 | 22.9 | 168308 | 2 | AC084775 Homo sapi |
| 41 | 474 | 22.2 | 87834 | 9 | AP002957 Homo sapi |
| 42 | 474 | 22.2 | 159946 | 2 | AC079194 Homo sapi |
| 43 | 458.5 | 21.5 | 93474 | 2 | AC136801 Rattus no |
| 44 | 458.5 | 21.5 | 173756 | 2 | AC134948 Rattus no |
| 45 | 458.5 | 21.5 | 245722 | 2 | AC120288 Rattus no |

ALIGNMENTS

RESULT 1


```

Db      1  ATGACGTGAAGAGAGAAAGCTTTACCGCTCGCTGACCCCGCGCCGCGAGCGAGCGC 60
Qy      21  ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db      61  CGGTACACCACTCGTCGCGGACAGCGAGGAGGCGCAAGCCCGCAGAAATCGTACAGC 120
Qy      41  SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db      121  TCCAGCGAGACCCCTGAAGGCTTAGCACCAGAGCGCCGCTAGCTATGAGCAGCGCGCTC 180
Qy      61  LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db      181  AAGACATTTGCGCGAGAGCGCGAGGAATCTGCCCGCAGGTGCCAATTCACCCCTG 240
Qy      81  ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db      241  CGGAGCTGGGCTGGGAAGATGAAGCCCTCAGCGGACCCCTGTACCGGACAGACATT 300
Qy      101  GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db      301  GGCCTCCCACTCGCGGTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 360
Qy      121  ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db      361  ACGTGTCTGTCCCTGAGCACCCCGTGGCTCTGTGGGCGCGAGACACCGTTCAGGGCGC 420
Qy      141  SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db      421  AGCTCTGTCTGTCCAGCGGGCCAAATTCACACTCACCGCACACCGGACATGAA 480
Qy      161  AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db      481  AACACTCAGACTGATCATCCGCGCGCGCTGCAGAACCAACCGCGCGCTCCGAGCCGCGC 540
Qy      181  ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db      541  CCGCGCTCTCGACAGCCACACCCCAACAGCACACCGCGGCTCCATTAATTCCTG 600
Qy      201  AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db      601  AACCGGGCAACTTCACGCGGAGAGCAACCCAGCGCGGCCCGCCACCGACCACTCGCTC 660
Qy      221  SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db      661  TCCGAGAGCCCTCTGCGCGCGCGCGCCAGAGCTCGCCACGCGCCCGAGGAATCGCTG 720
Qy      241  LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db      721  CTCACAGCAACATCCCGCTGGAGACCCAGAAACCTAGGCAAGCAGCCATTCTTAGGGACA 780
Qy      261  LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db      781  TTGAGGACAACTCATTTAGATGGACATTCGCGCGCTCCCGCCCATGATGGGGCTTAC 840
Qy      281  SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db      841  AGTACGGGACATTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTCTTCGCACCACATCA 900
Qy      301  ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db      901  CCAGGTACCCACTGAGCTGACGACAGTGTACTCTCTCGCGCCCGACCCCTCGCCCGC 960
Qy      321  SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db      961  AGCACTTCGCGCGCGCGCTTTAACTCAAGAGCGCTCCCAAGTACTGTAACTGGAAG 1020
Qy      341  CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db      1021  TGGCAGCCCTGAGCGGCATCGTCATCTCAGCCACTCTGGFCATCTCTGCTGGCATATTT 1080
Qy      361  ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

```

```

Db      1081  GTGCCCATGCACCTGTTTGGCTTAAACTGGCACCTGCAGCGATGGAGGGGAGATGTAT 1140
Qy      381  GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db      1141  GAGATCAGGAGGACACAGCAGCGAGTTGGCTGTGCCAACCGAGCTCTCCCTATACCCC 1200

RESULT 3
BD270887      3111 bp      DNA      linear      PAT 17-JUL-2003
LOCUS         Method of promoting the proliferation of inner ear hairy cells
DEFINITION   using ligand of HER2 receptor and/or HER3 receptor.
ACCESSION    BD270887
VERSION      BD270887.1 GI:33080655
KEYWORDS     JP 2002529425-A/6.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 3111)
AUTHORS      Gao, W.Q.
TITLE        Method of promoting the proliferation of inner ear hairy cells
JOURNAL      Patent: JP 2002529425-A 6 10-SEP-2002;
              GENENTECH INC
COMMENT      OS Homo sapiens (human)
              PN JP 2002529425-A/6
              PD 10-SEP-2002
              PF 28-OCT-1999 JP 2000580655
              PR 07-NOV-1998 US 60/107522
              PI WEI QIANG GAO
              PC A61K45/00,A61K38/00,A61K39/395,A61P27/16,C12N5/06//(C12N5/06,
              PC C12R1:91),
              PC A61K37/02,C12N5/00,(C12N5/00,C12R1:91)
              CC Method of promoting the proliferation of
              inner ear hairy cells
              CC using
              CC ligand of HER2 receptor and/or HER3 receptor
              FH Key Location/Qualifiers
              FT source 1..3111
              FT /organism='Homo sapiens (human)'
              source 1..3111
              /organism='Homo sapiens'
              /mol_type='genomic DNA'
              /db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 4,2e-103 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 6 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BD270887 (1-3111)
Qy      1  MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db      334  ATGACGTGAAGAGAGAGAAAGCTTACCGCTCGCTGACCCCGCGCGAGCGAGCGC 393
Qy      21  ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db      394  CGTACACCACTCGTCGCGGAGCAGCGAGGAGGCAAGCCCGCAGAAATCGTACAGC 453
Qy      41  SerSerGlnThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db      454  TCCAGCGAGACCTTGAAGGCTTAGCACCAGACCGCCGCTAGCTATGCGAGCGCGCTC 513
Qy      61  LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db      514  AAGGACATTTGCGCGAGGAGGCGGAGGAATTCGCGCGACAGGTGCCAATTCACCCCTG 573

```

| | | | |
|----|------|---|------|
| Qy | 81 | ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle | 100 |
| Db | 574 | CGGGAGCTGGGGCTGGAGAAGTAACAGCCCCCTCACGGGACCCCTGTACCGGACAGACATT | 633 |
| Qy | 101 | GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp | 120 |
| Db | 634 | GGCTCCCCCACTCGGGTACTCATCGGGGCTGGCTCTGATGCCGACATGGAGGCTGC | 693 |
| Qy | 121 | ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg | 140 |
| Db | 694 | ACGGTGTCTCCCTGTAGACACCCCGTGGCTCTGTGGGCGCGGACACACGGTCAGGGCGC | 753 |
| Qy | 141 | SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu | 160 |
| Db | 754 | AGTCTCTGCTGTCCAGCGGGGCCAAATTCCAATCTCACACTCACCGACACCGAGCATGAA | 813 |
| Qy | 161 | AsnThrGlnThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro | 180 |
| Db | 814 | AACACTGAGACTGATCATTCGGGGCGGCTGCAGAAACACCGCGCGCTCCGGACCGCGCGC | 873 |
| Qy | 181 | ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu | 200 |
| Db | 874 | CCGCGCTCTCGACGCCACACCCCCACACAGACACACCGCGGCTCTCATTAACCTCCCTG | 933 |
| Qy | 201 | AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu | 220 |
| Db | 934 | AACCGGGCAACTTTCAGCGGAGGAGCAACCCAGCGCGGCCCCACCGACCACTCGCTC | 993 |
| Qy | 221 | SerGlyGluProProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu | 240 |
| Db | 994 | TCCGGAGAGCCCTCTGCGCGCGCGCCCGAGAGGCTCGCCACGCCACGAGGAACTGGCTG | 1053 |
| Qy | 241 | LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr | 260 |
| Db | 1054 | CTCAACAGCAACATCCCCTGGAGACACAAACCTTAGCAAGCAGACCACTTCCTAGGACA | 1113 |
| Qy | 261 | LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr | 280 |
| Db | 1114 | TTGCAGGACAAACCTCATTTAGATGGACATTCCTCGGCGCTCCCGCCATGATGGGCTTAC | 1173 |
| Qy | 281 | SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer | 300 |
| Db | 1174 | AGTACCGGGCACTTCCTCTTCAAGCCCTGGAGGCACTCCCGCTCTTCCTGCACCACATCA | 1233 |
| Qy | 301 | ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg | 320 |
| Db | 1234 | CAAGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCTCGCCCCGACCCCTCGCCCCG | 1293 |
| Qy | 321 | SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys | 340 |
| Db | 1294 | AGCACCTTCGCGCGCGGCTTTAAACCTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG | 1353 |
| Qy | 341 | CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe | 360 |
| Db | 1354 | TGCGAGCGCTTGAGCGGCATCGTCATCTCACGCCACTCTGGTCATCTCTGCTGGCATCTTT | 1413 |
| Qy | 361 | ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr | 380 |
| Db | 1414 | GTGGCCATGCACCTGTTGGCTTAACCTGGCACCTGCAGCCGATGGAGGGGACATGTAT | 1473 |
| Qy | 381 | GluLeuThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro | 400 |
| Db | 1474 | GAGATCACGGAGGACACAGCCAGAGTTGGCTGTGCCAAACCGACGTCTCCCTATACCCC | 1533 |

RESULT 4

| | | | | | |
|------------|------------|---------------------------|-----|--------|-----------------|
| AR270934 | AR270934 | 3111 bp | DNA | linear | PAT 10-APR-2003 |
| LOCUS | Sequence | 1 from patent US 6500941. | | | |
| DEFINITION | AR270934 | | | | |
| ACCESSION | AR270934.1 | GI:29702176 | | | |
| VERSION | | | | | |
| KEYWORDS | | | | | |
| SOURCE | Unknown. | | | | |
| ORGANISM | Unknown. | | | | |

```
Qy 261 LeuGlnAspEnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGAGGACACACCTCATTTGAGATGGACATTTCTGGCGCCTCCCGCCATGATGGGGCTTAC 1173
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTCAGGGGACATTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTCTTGACACCATCA 1233
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGTATCCACCTGACGTCAGACACAGTGTACTCTCTCGCGCCGACCCCTGCCCGC 1293
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCGCGCGCGCTTTAACTCAAGAACGCGCTTCAAGTACTGTAACTGGAAG 1353
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCAGCCCTGAGCGCATCGTCATCTCAGCCACTCTGGTTCATCTCTGGCATACTTT 1413
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGCCCATGACACCTGTTGGCTTAACTGGCACCTGCAGCGATGGAGGGGACAGATGAT 1473
Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCGAGTGGCTGTGGCCTGTGCCACCGACGCTCTCCCTATACCC 1533

RESULT 5
BD140153 3111 bp DNA linear PAT 18-SEP-2002
LOCUS Utilization of heregulin as epithelial cell growth factor.
DEFINITION BD140153
ACCESSION BD140153
VERSION BD140153.1 GI:23235098
KEYWORDS JP 2002509076-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Sliwkowski,M. and Kern,J.A.
TITLE Utilization of heregulin as epithelial cell growth factor
JOURNAL Patent: JP 2002509076-A 6 26-MAR-2002;
GENENTECH INC, UNIVERSITY OF IOWA RESEARCH FOUNDATION
COMMENT OS Homo sapiens (human)
PN JP 2002509076-A/6
PD 26-MAR-2002
PF 03-FEB-1999 JP 2000530226
PR 04-FEB-1998 US 09/020598
PI MARK SLIWKOWSKI,JEFFREY A KERN
PC A61K38/00,A61K35/12,A61K39/395,A61K39/395,A61P11/00,A61P43/00,
PC C07K14/485,
PC C07K16/22,C12N15/09,C12P21/08,A61K37/02,C12N15/00 CC
Utilization of heregulin as epithelial cell growth factor PH Key
Location/Qualifiers
FT source 1..3111
FT /organism="Homo sapiens (human)".
FEATURES
source 1..3111
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.2e-103 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 6 Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x BD140153 (1-3111)
```

```
Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGACACGTGAAGGAGAGAAAGCCTTACCGCTCGCTGACCCCGCGCGAGCGAGCGC 393
Qy 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCACTCGTCCGCGGACAGGAGAGGCAAGCCCGCAGAAATCGGTACAGC 453
Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCTTGAAGGCTAGACACGAGCGCCCGCTAGCTATGGCAGCGCGTC 513
Qy 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 514 AAGGACATTTGGCGCAGGAGCGCGAGAAATTTTCGCGGCACAGGTGCCAATTCACCGCTG 573
Qy 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGGCTGGAGAAAGTAACGCCCTCACGGGACCCCTGTACCGACAGACATT 633
Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCGCACTGGCGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGSGCTGAC 693
Qy 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTGAGCACCCCGTGCCTCTGTGGGGCGCGAGCACACGGTCAGGGCGC 753
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTTCCTGCTGTCAGCGCGGCGCAATTCCAATCTCACCTCACCGACACCGAGCATGAA 813
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGGGCGCTCGACAGACCACCGCGGCTCCGAGCGCGCGC 873
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaSerIleAsnSerLeu 200
Db 874 CCGCGCTCTCGACGCGCCACACCCCAACAGCAGCACCGCGGCTCCATTAATCCCTG 933
Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCACGCGGAGAGCAACCCAGCGCGCGCCCGCCACCGACCATTCGCTC 993
Qy 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCCCTCTCGCGCGCGCGCCAGAGCTGCCACGCGCCAGGAGAACTGGCTG 1053
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCGCTGGAGACAGAAACCTAGGCAAGCAGCCATTCCTAGGGACA 1113
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGAGGACACCTCATTTGAGATGGACATTTCTCGCGCCTCCCGCCATGATGGGGCTTAC 1173
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTCAGGGGACATTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTCTTGACACCATCA 1233
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGTATCCACCTGACGTCAGACACAGTGTACTCTCTCGCGCCGACCCCTGCCCGC 1293
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCGCGCGCGCTTTAACTCAAGAACGCGCTTCAAGTACTGTAACTGGAAG 1353
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCAGCCCTGAGCGCATCGTTCATCTCAGCGACCTCTGGTTCATCTCTGGCATACTTT 1413
```

| | | | |
|---|----------|---|------|
| QY | 361 | ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr | 380 |
| Db | 1414 | GTGGCCATGCACCTGTTGGCTTAACCTGGACCTGCACCCGATGGAGGGGAGATGAT | 1473 |
| QY | 381 | GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro | 400 |
| Db | 1474 | GAGATCAGGAGGACACAGCCAGCAGTGGCTGTGGCAACCGACGCTCTCCCTATACCC | 1533 |
| RESULT 6 | | | |
| AF009227 | | 3111 bp mRNA linear PRI 17-SEP-1997 | |
| LOCUS | | Homo sapiens gamma-heretulin mRNA, complete cds. | |
| DEFINITION | | AF009227 | |
| ACCESSION | | AF009227 | |
| VERSION | | AF009227.1 GI:2406643 | |
| KEYWORDS | | | |
| SOURCE | | Homo sapiens (human) | |
| ORGANISM | | Homo sapiens | |
| REFERENCE | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| AUTHORS | | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| TITLE | | 1 (bases 1 to 3111) | |
| JOURNAL | | Schaefer, G., Fitzpatrick, V.D. and Sliwkowski, M.X. | |
| MEDLINE | | Gamma-heretulin: a novel heregulin isoform that is an autocrine | |
| PUBMED | | growth factor for the human breast cancer cell line, MDA-MB-175 | |
| AUTHORS | | Oncogene 15 (12), 1385-1394 (1997) | |
| TITLE | | 97472144 | |
| JOURNAL | | 9333014 | |
| MEDLINE | | 2 (bases 1 to 3111) | |
| PUBMED | | Schaefer, G., Fitzpatrick, V.D. and Sliwkowski, M.X. | |
| AUTHORS | | Direct Submission | |
| TITLE | | Submitted (18-JUN-1997) Protein Chemistry, Genentech, Inc., 460 | |
| JOURNAL | | Point San Bruno Blvd, South San Francisco, CA 94080, USA | |
| FEATURES | | Location/Qualifiers | |
| Source | | 1..3111 | |
| | | /organism="Homo sapiens" | |
| | | /mol_type="mRNA" | |
| | | /db_xref="taxon:9606" | |
| | | /chromosomes="8" | |
| | | /map="8p22-p11" | |
| | | /cell_line="MDA-MB-175" | |
| | | /cell_type="breast cancer cell" | |
| | | 334..2640 | |
| | | /codon_start=1 | |
| | | /product="gamma-heretulin" | |
| | | /protein_id="AAC51756.1" | |
| | | /db_xref="GI:2406644" | |
| | | /translation="MDVKEKPYRSLTRRDAERYTSSADSEKAPQKYSSET | |
| | | LKAYDDARLATGSRVKDILVPOEAEFCRTGANFTIRELGLSEVTPPHGTLRTDGL | |
| | | PHCYGSMGAGSDADMEADTVLSPHPVRLWGRSTRSGSSCLSSRANSNLITLDTPEH | |
| | | NTEIDPGQLQNHARLTPPPPLSHAHTPNQHAASINSLRNFTPRSNPSPATDH | |
| | | SLSGPPAGAQEPAHAQBNLLNSNIPLETNLGKQPLGLTQNLIEBMDILGASRH | |
| | | DCAYSDGHFLFKPGGTSPLCFTTSPGYPLTSTVSPPLPRSTFAPAPNLAKKPS | |
| | | KYCNWKAALSAIVISATLIVILLAVFVAMHFLGNWHLOPMESQMVLTEDTASWPV | |
| | | PDVSLYSGGTGLTPDRKKGHTTEKSPSPEDSFIDSGEIDVGRASOKIPPT | |
| | | FWRQVFDHPVHLKFNPSLGAALVGIYGRKGLPPSHFTQDFVELLDGRRLLTQEAR | |
| | | LEMTPRQSGRTVPSPSHETGFIQYLDGSHLAFYNDGKESEVSLFTALPPRL | |
| | | SKMGQSAAGSLKVLRCSTSEYSLRFKFNKGNELNKNKPNQIKQKPKPSL | |
| | | RINKASLDGSMCKVISKLGNDSANITIVESNEIITGMPASTEGAYVSESPIR | |
| | | LSVSTEGANTSSTSTSTTGTGLVKVKAKEKTEPCVNGGCECFWVLDLSNPRLCKP | |
| | | NEFTGDRCONYMASFYSTSTPFLSLPE" | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 4.2e-103 | Length: | 3111 |
| Score: | 2122.00 | Matches: | 396 |
| Percent Similarity: | 99.00% | Conservative: | 0 |
| Best Local Similarity: | 99.00% | Mismatches: | 4 |
| Query Match: | 99.53% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |
| SEQ14-X-AT-28-64-76 (1-400) x AF009227 (1-3111) | | | |
| QY | 1 | MetAspValLysGluAlaGlyProTyrArgSerLeuThrArgArgAspAlaGluArg | 20 |
| Db | | | |

| | | | |
|----|------|---|------|
| Db | 334 | ATGACGCTGAGGAGGAGGAGCGCTTACCGCTCGCTGACCCCGCGCGCGCGCGCGC | 393 |
| QY | 21 | ArgTyrThrSerSerSerAlaA***SerGluGluGlyLysAlaProGlnLysSerTyrSer | 40 |
| Db | 394 | CGCTACACCGAGCTCTGCGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 453 |
| QY | 41 | SerSerGluThrLeuLysAlaTyrAspGlnAspAlaAArgLeuAlaTyrGlySerArgVal | 60 |
| Db | 454 | TCCAGGAGAGACCTGAGGCTTACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG | 513 |
| QY | 61 | LysAspLys***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu | 80 |
| Db | 514 | AAGGACATTTCTCGCAGGAGGCGGAGGAAATCTGCGGCACAGGAGGAGGAGGAG | 573 |
| QY | 81 | ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspLys | 100 |
| Db | 574 | CGGAGAGTGGGGCTGGAAGAGTAACGCCCTTACGGGACCTGTACCGGACAGACATT | 633 |
| QY | 101 | GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp | 120 |
| Db | 634 | GGCTCCCGCCACTGCGGCTACTCCATGGGGGCTGGCTCTGTATGCCGACATGGAGGCTGAC | 693 |
| QY | 121 | ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg | 140 |
| Db | 694 | ACGCTGCTGCTGCGCTGAGCACCCCGTCTCTGTGGGGCGGAGCACACGCTCAGGCGC | 753 |
| QY | 141 | SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu | 160 |
| Db | 754 | AGCTCTGCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG | 813 |
| QY | 161 | AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro | 180 |
| Db | 814 | AACTCTGAGACTGATCATCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG | 873 |
| QY | 181 | ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaSerLeuAsnSerLeu | 200 |
| Db | 874 | CGCGCGCTCTCGCACGCCACACCCCGCCAAACAGCAGCACCGCGCGGCTTAACTCCCTG | 933 |
| QY | 201 | AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu | 220 |
| Db | 934 | AAACCGGGGCACTTACGCGCGAGGAGCAACCCCGCGCGCGCGCGCGCGCGCGCGCTC | 993 |
| QY | 221 | SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu | 240 |
| Db | 994 | TCCGAGAGAGCCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG | 1053 |
| QY | 241 | LeuAsnSerAsnLleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr | 260 |
| Db | 1054 | CTCAACAGCAACATCCCTCGGAGACCCAGAAACCTAGGCAAGCAGCCATTCTCTAGGGACA | 1113 |
| QY | 261 | LeuGlnAspAsnLeuLeuMetAspLysLeuGlyAlaSerArgHisAspGlyAlaTyr | 280 |
| Db | 1114 | TTGAGGACAACTCATTTGAGATGACATCTCTCGCGCGCTCCCGCATCATGGGGCTTAC | 1173 |
| QY | 281 | SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer | 300 |
| Db | 1174 | AGTGACGGCACTTCTCTTCAAGCCTGAGGAGCACCTCCCGCGCTCTCTCTGCACCATCA | 1233 |
| QY | 301 | ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg | 320 |
| Db | 1234 | CCAGGGTACCCACTACGTCACGACAGGTGATCTCTCTCCCGCGCGCGCGCGCGCGC | 1293 |
| QY | 321 | SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys | 340 |
| Db | 1294 | AGCACCTTCGCGCGCGCGCGCTTTAACTCAAGAGAGCCCTCCCAAGTACTGTAATCGAAG | 1353 |
| QY | 341 | CysAlaLeuSerAlaLleValleSerAlaThrLeuValleLeuLeuAlaTyrPhe | 360 |
| Db | 1354 | TCCGAGAGCCCTGAGCGCCATCTCTCAGGACCTCTCTGTCATCTCTGTCGATCTTT | 1413 |
| QY | 361 | ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr | 380 |
| Db | 1414 | GTGGCCATGCACCTGTTGGCGCTAACTGCGACCTGCACCGCGGAGGAGGAGATGAT | 1473 |

```

QY 381 GluileThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCAGTGGCTGTGCCAACCGACGTCCTCCATATACCCC 1533

RESULT 7
AX675551
LOCUS AX675551 8438 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 1 from Patent WO02055704.
ACCESSION AX675551
VERSION AX675551.1 GI:29333552
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Padigaru, M., Li, L., Zerhusen, B.D., Casman, S.J., Shenoy, S.,
  Spytek, K.A., Zhong, M., Gangolli, E.A., Burgess, C.E., Patturajan, M.,
  Vernet, C.A., Taylor, S., Tchernev, V.T., Miller, C.E., Guo, X.,
  Boldog, F.L., Grosse, W.M., Alsobrook, J.P., Gerlach, V.,
  Edingermark, S., Rothenberg, M.E., Ellerman, K., Macdougall, J.,
  Malyankar, U., Millet, I., Peyman, J., Smithson, G., Gunther, E. and
  Stone, D.J.
TITLE Proteins, polynucleotides encoding them and methods of using the
same
JOURNAL Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)
FEATURES
source Location/Qualifiers
1..8438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.6e-100 Length: 8438
Score: 2082.00 Matches: 395
Percent Similarity: 98.01% Conservative: 0
Best Local Similarity: 98.01% Mismatches: 5
Query Match: 97.65% Indels: 3
DB: 6 Gaps: 1

SEQ14-X-AT-28-64-76 (1-400) x AX675551 (1-8438)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 4 ATGACGTGAAGGAGAGAAAGCCCTTACCGCTCGCTGACCCCGCGCCGACGCGAGCGC 63

QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 64 CGTACACACAGCTCGTCCGCGGACGAGGAGGCGAAGCCCGCAGAAATCGTACAGC 123

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 124 TCAGCGAGAGACCCCTGAAGGCGCTACGACAGCAGCGCGCGCTAGCCTATGCGAGCGCGTC 183

QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 184 AAGGACATTGTGCGGAGGAGGCGGAGAAATCTGCGCGACAGGTGCGCAACTTCACCCCTG 243

QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 244 CGGAGCTGGGGCTGGGAAGAGTAACCGCCCTCCACGGGACCTGTACCGGACAGACATT 303

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 304 GGCTCTCCCACTGCGGCTACTCATGGGGCTGGCTCTCATGCGGACATGGAGGCTGAC 363

QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 364 ACGGTGCTGTCCTGAGCACCCCGCTGCTGTGGGGCCGAGCACACGTCAGGCGCGC 423

```

```

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 424 AGCTCTCGCTGTCCAGCGGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 483

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 484 AACACTGAGACTGATCATCCGGCGGCTGTGAGAACCAACCGCGGCTCCGAGCGCGCG 543

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAsnSerIleAsnSerLeu 200
Db 544 CGCGGCTCTCGAGCGCCACACCCCAACAGCAGCAGCGGCTCCATTAACTTCCCTG 603

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 604 AACCGGGCAACTTTCACCGCGAGGAGCAACCCAGCGCGGCCCGCCACCGACCTCGCTC 663

QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 664 TCCGAGAGCCCTTCCCGCGCGGCGCCAGAGGCTGCCACAGCCCGAGAGAACTGGCTG 723

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 724 CTCACAGCAACATCCCTCGAGCAGCAAACTTAGCAGCAGCAGCACTTCTTAGGGACA 783

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 784 TTGAGGACAACTCATTTAGATGACATTCCTCGCGCTCCCGCATGATGGGCTTAC 843

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 844 AGTACCGGCACTTCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTTGACACCATCA 903

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 904 CAGGGTACCCACTGACGTCCAGCACAGTGACTCTCTCGCGCCCGACCCCTGCCCGC 963

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
Db 964 AGCACCTTCGCGTGGCGGCTTTTAACTCAAGAGGCGCTCCAAAGTACTGTAACTGGAAG 1023

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1024 TGGCGACCTTGAGCGCATCGTCATCTCAGCCACTCTGGTCACTCTGCTGGCATATTT 1083

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet--- 379
Db 1084 GTGGCATGCACCTGTTGGCCTAAACTGGCACCTGCAGCGGATGGAGGGGCGAGATGACG 1143

QY 380 ---TyrGluileThrGluAspThrAlaSerSerTrpProValProThrAspValSerLe 398
Db 1144 GATTTATGATCAGCGAGGACACAGCCAGTGGCTGTGCCACTGTGCCAACCGACGTCCTC 1203

QY 398 uTyrPro 400
Db 1204 ATACCC 1210

RESULT 8
AX600210 8645 bp DNA linear PAT 14-FEB-2003
LOCUS Sequence 22 from Patent WO02072830.
DEFINITION AX600210
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Yue, H., Yao, M.G., Ison, C.H., Lu, Y., Warren, B.A., Elliott, V.S.,
  Baughn, M.R., Ding, L., Xu, Y., Gietzen, K.J., Tang, T.Y., Lal, P.G.,
  Duggan, B.M., Burford, N., Lu, D.A., Richardson, T.W., Tran, U.K.,
  Khare, R. and Wallia, N.K.
TITLE Proteins associated with cell growth, differentiation, and death

```

JOURNAL Patent: WO 02072930-A 22 19-SEP-2002;

Incyte Genomics, Inc. (US)

Location/Qualifiers

FEATURES

source

1..8645

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

/note="Incyte ID No: 7489573CB1"

ORIGIN

Alignment Scores:

Pred. No.: 1,14e-99 Length: 8645

Score: 2066.00 Matches: 390

Percent Similarity: 97.50% Conservative: 0

Best Local Similarity: 97.50% Mismatches: 6

Query Match: 96.90% Indels: 4

DB: 6 Gaps: 2

SEQ14-X-AT-28-64-76 (1-400) x AX600210 (1-8645)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
 DB 117 ATGGACCTGAAGAGAGAGAGCTTACCGCTCGCTGACCGCGCGCGCGAGCGC 176
 QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
 DB 177 CGCTACACAGCTCGTCCCGGAGACGAGAGAGGCGAAGCCCGCAGAAATCGTACAGC 236
 QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
 DB 237 TCCAGCGAGACCTGAAGCGCTACGACGAGGACCGCGCGCTAGCTATGGCAGCGCGTC 296
 QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
 DB 297 AAGGACATTTGTCGCGAGGAGCGGAGGAATTCGCGCACAGGTCCCACTTCACCCCTG 356
 QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
 DB 357 CGGAGCTGGGGTGGAGAGAGTAAAGCCCTCACGGGACCTGTACCGGACAGACATT 416
 QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
 DB 417 GGCCTCCCCACTCGCGTACTCATCGGGGCTGGCTCTGATCCGACATGGAGGTGAC 476
 QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
 DB 477 ACGGTCTGTCTCCCTGAGCACCCCGTCTGTGGGCGCGAGCACACGTCAGGCGC 536
 QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
 DB 537 AGCTCTGCTGTCCAGCGCGGCGCAATTCATCTCACATCCCGACACCGAGCATGAA 596
 QY 161 AsnThrGluThrAspHisProGlyGlyLeuGluAsnHisAlaArgLeuArgThrProPro 180
 DB 597 AACACTGAGACT-----CCGGCGCGCTGCAGAACCCAGCGCGGCTCCGGAGCGCGCG 650
 QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
 DB 651 CCGCGCGCTCTCGACGCGCGCACCCCCAACAGCACCCCGCGGCTCCATTAACTCCCTG 710
 QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
 DB 711 AACCGGCGCACTTCAGCGCGAGAGAGCAACCCAGCGCGCGCGCGCGCGCACTCGCTC 770
 QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnThrLeu 240
 DB 771 TCCGAGAGCGCCCTGCG 830
 QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
 DB 831 CTCACAGCACATCCCCCTCGAGGACGCAAACTTAGCAGCAGCATCTCTTAGGACA 890
 QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280

DB 891 TTGCAGGACAACCTCATTGAGATGACATTTCTGGCGGCTCTCCGCCATGATGGGCTTAC 950
 QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
 DB 951 AGTACGGGCGCACTTCTTCAAGCCTGGAGCACCTCCCGCTCTTTCGACCAATCA 1010
 QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
 DB 1011 CCAGGGTACCCACTGACGCTCCAGCACAGTACTCTCTCCGCCCGGACCCCTGCCCGC 1070
 QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnThrLys 340
 DB 1071 AGCACCTTCGCCCGCGCGCTTTAACTCAAGAGAGCCCTCCCAAGTACTGTAACTGGAAG 1130
 QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
 DB 1131 TCGCGAGCCCTGAGCGCCATCTCATCTCAGCACCTCTCGTCTCTCTGCGCATCTTT 1190
 QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
 DB 1191 GTGGGTAAACACCTCTTC-----AACTGGCACTTGACCGGATGGAGGCGAGATGAT 1244
 QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
 DB 1245 GAGATCAGCGAGACACAGCCAGCAGTTGGCCTGTGCCAACCCAGCTCTCTCTATACCC 1304

RESULT 9

AB025413

LOCUS

DEFINITION

AB025413

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..8585

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Balb/c"

/db_xref="taxon:10090"

/tissue_type="brain"

/dev_stage="adult"

1..8585

/gene="ten-m4"

188..8503

/genes="ten-m4"

/codon_start=1

/product="Ten-m4"

/protein_id="BAA77399.1"

/db_xref="GI:4760782"

/translation="MDVKMKPYRSLTRRDAERRYTSSADSEEGKPKQKSYSSSET

LKAYDDARLAYSGRMDKMPVPAEFCRTGNTFTRELGLGEMTPPHGIDYRDI

PHCGSYNGASSDADTADTDLSPHVPVLRSGRSLCSLRASNTLTITDEH

NTTDPSSLQNHPLRTPPPLPHATPNQHAASINSLNRGNTFRSNPSPATDH

SLSEPPAGSAQEPHQAQNDNLLNSNIPLETNRNLKQKPFILGTLDNLIEMDILSASRH

1..8585

/gene="ten-m4"

188..8503

/genes="ten-m4"

/codon_start=1

/product="Ten-m4"

/protein_id="BAA77399.1"

/db_xref="GI:4760782"

/translation="MDVKMKPYRSLTRRDAERRYTSSADSEEGKPKQKSYSSSET

KEYWORDS
SOURCE
ORGANISM

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 9722)
Wang, X.Z., Kuroda, M., Sok, J., Batchvarova, N., Kimmel, R., Chung, P.,
Zinszner, H. and Ron, D.
Identification of novel stress-induced genes downstream of chop
EMBO J. 17 (13), 3619-3630 (1998)

MEDLINE

98315054

PUBMED

9649432

REFERENCE

2 (bases 1 to 9722)

Wang, X.-Z. and Ron, D.

AUTHORS

TITLE

JOURNAL

Location/Qualifiers

1..9722

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NIH/3T3"

/db_xref="taxon:10090"

/cell_line="NIH-3T3"

/note="tunicamycin-treated"

1..9722

/genes="Doc4"

/note="downstream of CHOP 4"

583..9060

/genes="Doc4"

/note="similar to Drosophila melanogaster temm/odz and

human gamma-heretulin; type II transmembrane protein"

/codon_start=1

/product="Doc4"

/protein_id="AAC31807.1"

/db_xref="GI:3170615"

/translations="MDVKRKYRSLRRDAERRVTSSEADSEKGGPKSYSSSET

LKAYDQRLAYGSRVQKDVPOQAEFRCTGTNTRELGGLGEMTPPHGTLXTDGL

PHGYSGLASSADLEADVLSPEHVRILMGSTRGRSCLSSRANSNLITLDTDEH

NTETGAPLCHSSASTPIEQSPFPSPANESORILLGNVQAQPTDSEEFVFN

SLVKGSGASLGAANDHPSLQNLHRLTPPLPHAPTHQHAASINSLNKNGFT

PRNPSPADHLSGEPAGSAQEPHTAQDNVLSKIPVETRNKLGKQPLFTWQDN

LIMDIPFASRRGQSDGHHFFPKGTSPLFTSPGPLTSTVSPRPPLRST

FSPAPNLKSPKYNKCAALSAILISATLIVLLAVFVAMHLFGLNHLQPMWQMO

MYEITDTSASVPTDVSILYPSGGTGLETPDRKGAEGKPSLFPEDSFIDSGEI

DVGRASQKIPPTGFRSQRVFIIDPHVLKENVSLGKAALVIGYGRKGLPSHSTQLQFV

ELDGRRLTLEARSLEGRQRPVPPSSHETGFTQYLDSDGIVLHAFYNDKSESEV

VSLTITAIESVDNCPNCGNDCISGTCFLGFLGDCGRASCPFLVCSNGQYMKG

RCLCHGWKGAECVPTNQCIDVACSHGTCTIMGTICINPGYKGESCEYVDMDPTCS

SRGVCGECHCSVGMWGTCTETPRATCDQCSGHGTFLPDTGLCNDPQSWTGHDCSI

EICAADCGTSGVCGTCTCEDGWMGAACDQACHPCABSHGTCDGCKRCTPGWNE

HCTIEGCPGNGRCTLDELNGHVCVQLGWRGTGCDTSMETGCGDKNDGDLVD

CMDDPCLQPLCHVNPCLGSPDPLDIIQETAPVSOQNLNSFYDRIKFLVGRDTHS

IPGENPDGHACVIRQVMTSDGTLVGVNIIFINPLFGTISRQDGSFVLVINGG

ISILIRPERAPFTTQETHLWLPMDRPFVMTIVMRHEEIPSRDLNFAFPNFPVSP

SPLTSPASSCAEGPIVPEIQALQEEIVIAGCKRLSYLSRTPGYKSVVRISLTHPT

IPNLNMVLMVAVEGLFRKFAAPDLVYFLWDKTVYVNOVKVGLSEAFVSGVE

YESCPDLILKEKRTAVLQGEIDASKLGGSLDKHNLNTQSLILHKGNGENFVSQO

PVVGITMNGRRRSICSPCNGLADGNKLLAPALTCGSDGSLVGFYFNIRIRFES

GNVTNILEMHSFAKHYLYLATPMGSAFLSDTNSRVFKVSTTVVVDLVKNSVVA

GTGDCIPLFDPTDTCGGKATEATLNPRTGTVDKFGLIYFVDGTMIRVDQNGILST

LGNSNDLTSRPLSCGSMBISQVRLWPBTDLAINPMDNSLYLDNNDVQLISENHQV

RVAGRPMHCVPGIDQFLLSKVAIHATLESATALAVSHNGVLIYAETDEKKINRIQ

VTTSGELSLVAGPSGCDKNDACDPSGDDGYAKDAKLTNPSSLVACADGELYVAD

LGNIIRIFIRIIRTNQNYELSSPDLQELVLPDTSIGKHLIYQSLPTGDLYLNFY

TGDDITHIRIDNNGNMNVNRDSTMPLWLVPDQVYVMTWTFNSALRSVITQGH

AMTYHNGSGLLATKSNENGTTFEYDVSRLNTVTFPTQVNSFSRSDTSSVHVQV

ETSSKDDVITLHLSGSAFTLQDQVRSYICADGSLFLLLANGNEVALQTEPHL

LACTVNTVGRNVTLPIDGLNLVEMRQKQGVTVPRPLRVNRLNLSLDFD

RVETEKIDHRRFTLRLYDQAGRSFWSPSRLNGVNVTVSPGGHIAIGIQRGIMS

ERMYDQAGRTLSRTIFADGKMSYTYLEKSNVLHLHSQRQVIFEDFNKDRLLSSVTMN

VARQTLTISVGYRNIYQPEGNASVYQDFTEDGHLHLLTFLYLGTRRVLRYTKGLS

KLAETLYDITTKVSTYDTAGMLKTVNLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGM

misc_feature

1813..1881

/gene="Doc4"

/note="putative; transmembrane-region site"

ORIGIN

Alignment Scores:

Pred. No.: 5.68e-94 Length: 9722

Score: 1959.00 Matches: 373

Percent Similarity: 80.72% Conservative: 8

Best Local Similarity: 79.03% Mismatches: 19

Query Match: 91.89% Indels: 72

DB: 10 Gaps: 2

SEQ14-X-AT-28-64-76 (1-400) x AF059485 (1-9722)

QY 1 MetAspValIysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20

Db 583 ATGGAGCGTGAAGGAGAGAGCCCTATCGCTCGCTGACCCGGCGTGCAGTGTGAGCGC 642

QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40

Db 643 CGCTACACGAGCTCATCGGACAGGAGGAGGCGCAAGGGCCACACAGAAAGTCTACAGC 702

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuLysAlaTyrGlySerArgVal 60

Db 703 TCCAGTGAGACCTGAAGGCATATGACCAAGATGCTCGCTGACCTACGCGACGCGGCTC 762

QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80

Db 763 AAGGACATGTATACACAGGAGCGGAGGATCTCGCGCACACAGCACACTAATTTCACTCTT 822

QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100

Db 823 CGTGAACCTGGAGCTGGGAGAGATGACGCCGCCCTCATGGGACTCTCTACAGGACAGACATC 882

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120

Db 883 GGCTTCCACACACTGTGGCTATTCCATGGGGGCGAGCTCTCATGCGAGACCTGGAAAGCAGAC 942

QY 121 ThrValLeuSerProGluHisProValArgLeuTyrArgSerThrArgSerGlyArg 140

Db 943 ACTGTGCTATCCCTTGACACCCAGTGGCGCTGTGGGGCGGAGACACACGCTCAGGCGCGC 1002

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160

Db 1003 AGTTCTCTGCTGTCAGCGCGGCGCAACTCCAAACCTCACTACCGGACACAGACGATGAG 1062

QY 161 AsnThrGluThr 164

Db 1063 AACACAAACCGGTGTCTCCCTTGCAATTGTCATCTCTGCTTCATCAACCCCTATTGAGCAG 1122

QY 164 164

Db 1123 TCCCCCTCCCGCCCCCTCCCTCCCGCAATGAGAGCCAGAGCGGCTTGTAGGCAAC 1182

QY 164 164

Db 1183 GGTGTGGCCACAGCAACCCCGGACTCAGACTCTGAGGAAGAGTTTGTCCCTAAATTCATTC 1242

QY 165 165

AspHisProGlyGlyLeu 170

Db 1243 TTAGTAAAGTGGCTCGCCAGCGCTGGGGTCCGAGCAACGATCATCCAGAGCCTG 1302
Qy 171 GlnAsnHisAlaAArgLeuArgThrProProProLeuSerHisAlaHisThrProAsn 190
Db 1303 CAGAACCCCTCGGCTCGGACGGCTCTCCACCACTGCCCGACGCCCATACCCCCAC 1362
Qy 191 GlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn 210
Db 1363 CAGCATACCGGGCTCATCACTCTTGAACAGGGCACTTCAACCCAGGAGCAAC 1422
Qy 211 ProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln 230
Db 1423 CCCAGCCAGCACCCACAGACCACTCACTCTCTGGGAGCCCCAGCAGCAGCCAG 1482
Qy 231 GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArg 250
Db 1483 GAGCCAAACCATGCCAGACCACTGGGTACTCAACAGTAAATCCAGTGAACACCA 1542
Qy 251 AsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIle 270
Db 1543 AACCTAGGCAAGCAGCATCTCTAGGACATGGCAGGACCACTCATTTAGATGACATT 1602
Qy 271 LeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysProGly 290
Db 1603 TTCAGCGCTCCCGCGCTGATGGGGCTTACAGTGACGGGCACTTCTTTCAGCCCGA 1662
Qy 291 GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrVal 310
Db 1663 GGCACCTCCCGCTCTCTCTGCACACATCCCGAGGGTACCCCGCTAACCTCTAGCAC 1722
Qy 311 TyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeu 330
Db 1723 TATTCGCCCCACCCCGCCCTGCCCCGAGCACTTCTCCGACGACCTTTTAACTC 1782
Qy 331 LysLysProSerLysTyrCysAsnTrpLysCysAlaLeuSerAlaIleValIleSer 350
Db 1783 AAGAAGCCTTCAAGTACTGCAACTGGAAGTGTGAGCGCTGAGCGCCATCTCATCTCA 1842
Qy 351 AlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrp 370
Db 1843 GCTACGCTGTCATCTCTCGCATCTTGTGGCCATGCACTTGTGGCTTAACCTGG 1902
Qy 371 HisLeuGlnProMetGluGly-----GlnMetTyrGluIleThrGluAspThrAlaSer 388
Db 1903 CACCTGAGCGGATGAGGGGCGAGATGAGATGATGAGATCAAGGAGACACAGCCAGC 1962
Qy 389 SerTrpProValProThrAspValSerLeuTyrPro 400
Db 1963 AGCTGGCTGTGCCAACGAGCTCTCCCTGTATCCC 1998

RESULT 11

AB026980
LOCUS AB026980 9264 bp mRNA linear VRT 05-NOV-1999
DEFINITION Danio rerio mRNA for ten-m4, complete cds.
ACCESSION AB026980
VERSION AB026980.1 GI:5307784
KEYWORDS ten-m4.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (sites)
AUTHORS Mieda, M., Kikuchi, Y., Hirate, Y., Aoki, M. and Okamoto, H.
TITLE Compartmentalized expression of zebrafish ten-m3 and ten-m4,
homologues of the Drosophila ten(m)/odd Oz gene, in the central
nervous system
JOURNAL Mech. Dev. 87 (1-2), 223-227 (1999)
MEDLINE 99425191
PUBMED 10495292
REFERENCE 2 (bases 1 to 9264)
AUTHORS Mieda, M.
TITLE Direct Submission

JOURNAL

Submitted (06-MAY-1999) Michihiro Mieda, Brain Science Institute,
RIKEN, Lab. for Developmental Gene Regulation; 2-1 Hirosawa,
Wako-shi, Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp,
Tel:81-48-467-9713, Fax:81-48-467-9714)
Sequence updated (29-Jun-1999).

COMMENT

FEATURES

source

1..9264
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/dev_stage="embryo"
1..9264
/gene="ten-m4"
323..8797
/note="similar to Drosophila melanogaster tenm/odz and Mus
musculus Doc4"
/codon_start=1
/product="ten-m4"
/protein_id="BAA81893.1"
/db_xref="GI:5307785"

gene

323..8797

CDS

/translation="MEVKRRPRLSLTSSRRTRTSSADSDGKINPKSYSSSET
LKADQDRLAYGSRVKDLVHHEDEPSRQGPDLRDMARQDPVPHMGAYRTEMGL
PHRYSVSADADTETDGTMSPEHAYRLGRSNTKGRSSCLFSRANSNLTIDTEH
ENTNGPLHCSSASSSFDPFPPSHAAQSQRLIGNSGAGRDESEDFBGEN
SLVKTSGNVTAPAAATANEFSQNSHRLTPEPLPFHSHSPQHHTASIGLSRSN
YTKNSPSPAPTDSSANEGETSAQSSAQNMLNSNVLPTNRATKOTFLETLQD
NFIEMDILATARRDCAVTDGHFLFKPGTSPLYCTSPGTVSPVPPRLPRN
TFSPATSLKPKYKHCNWKCAALISLTVLFLAVFIAMHLFGLNHLQVQVQRI
YQLTNTSGHLPLDLGLPLGNTGLFEPRGSRDGLKDFPPEISFLDMEIDVG
RKVAQLIPGFIWRSQVFIHPIWMLKFNVSLSKDALVTIGRRGLPSPHSHTQFDEVELL
DGRLLSGLDGLDGPFPFPAQRLVITSHDTGCIQYMDSGIWHLAVINDGKRETV
SFETATLSDIDDCNSGNGDCVGNCHFCPPGPRGDCSRASCPVLCSNGOYLKGR
CMCHSGKSGECVPTNQCIDITCSGHGTCTGTCICNPYKGNCEVDCDLPKCSG
RGVCRCEHCFCVWGPGGCPSPASCMEOCSGHGSFLADTNTCNCHNTHDQCSIE
LCAADCGHGI CVAGSCRCDEGMWTCGEQACHPCRSSEHGTCKDKCESCPWNGEH
CTIEGCPGLNGNRCITLNGNGYCVCLGWRGAGCDTSMETACSGDRNDGDLGIDTC
MDPCCQAQCHTSLTSCVSPDLIIIOETISSLSTLSQSYQRIHFLVRGDSHTVI
PDVNPFDGIHACVIRGQVVTSDGTPLVGNISFINKPAYGTYITRQSGSFDLVNGGV
AIGLRFERAPFTIQTETLWLPGRFFVMDTIVMRHEVNDIPSCDLSFTPMPLVLA
PLTAFACCTGPERGVVPEIOTLOEVRIPGTDMLGLYLSRTSGYKSLRLTLTSTI
PFSLMKVHLMAVEGRLFRKWSAAPNLISDFWDXTDVYSQYKYLSEAFVSQFEY
ESCPDILLWEKRTAVLQGYETTASNGSVDRHHLNIOGILHNGENIISQOP
PVTSGMNGRRRSISCPNGLADGNKLLAPALACSGDSLGVGDFNVYRRTFTTG
NVTSLSLSPAKHYLATSPVSGMLYLSDTSSRKVKVKSLYAVDKVAKNLBVLG
TGQCLPYDTRDGKGAVENTLTPRGITVDKYGVIFFDVTMIRRIDQNGIISTL
LGFNDLTSARPLSCDSVMDISOVRLEWPTDLAVSPMNSLYLVLDNNVLOI SENHOVR
IVAGRPMHCOVPLGLDFVSKIAIHATLESANALAVSHNGLIYIAESDEKINRVQV
STNGEISLLAGAPSGCDKNDACDCYSGDGYAKAKLAPSLAVSPDGEFLADL
GNIRYVRNKAFLNPLNMYEISSPIDDELYLFDVNASHVFTQSITGDIYLYNTYS
EGDLSITDKNNRVSIRRDSTGLPLMLMGPDQTFWTMTGNALSKSAAQOEIA
VMYHSGSLLATKSNEDGWSFTFYDNYGRLNTVYPTGRVSRYRDSYTRVQTE
GSKNEDITVTNLISAGTFTYLMQDVKNYSYIIGLGLRLVLVLANGNEVSLHTPEHL
SGTNPTISKRNVTLPIDNGLNVQVRKEQAGQVTVYGRRLVRHNRLISMDPR
VTRTEKYDDHDKETRLRIHVDHAGRPILWAPSRNLGNVNTYSPGCHAGIQRGTMSV
RMEYDQNGRTISKI PADKWSYTYLEKSWLLYSQRYIFEDKNDRLSYSTMPNV
ARQLETTRISGIYRNTYRPEGNATVQYSEDGLLQTHOQTGRVYIKYKGLSR
LLEILYDTRIAFYSDESAGMLKTVGLQSEGFACITRYRQIGLQIFRISSEGMV
NARFDYNDNSFRVTSQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYVDNQIIT
TAYWTHPKHFDAYGRVQVOYEI PRSLMYMMVQFONMGRVAVAKLVGVPYNTRYA
VEYDADQLOVSDINQKPLWYSYIDLNGNLHLLSPGNSARLTPLRYDIRIRITLGDV
QYRLDEDFLRQNGDNFFENYAGLLVNTKVNQWTIKYRYDGLGRRVSRSTQGH
LQFFYADLSIPRTHMYNHSSEITSLYYDGLHFLAMELSSGDEFFVACDNIGTPL
AFVSGAGMLIKQILTHAFGEYLDNSPFLVIGYQGLYEPILTKLVHMRDYLVA
GRWTTDPDHIRKLDNSDINPVNFKNNNSQETKCYMTDMSLWLTGVYRFGOLY
NVIPYKRPVTDAMEPSYELVHTQIKTQEWDSKSVLGVQCEVORQLSKFVRLERQ
IYSASGCGCTPLHLTATGTSIFGKGVKVAIREGVERADIISLANEDGRRITAAVD
KASYLQDLHLTITAGTLITFVKSGVEGLSLGTMVQGTLETGNTVSNVQYNNMVLG
GSRRLTDIQMYGTLSLNVGSSVEKRVLELQRAVATAWAHIERLRQEE
GSRWTDGERQQLLSRGVQGYEGFYIVSDQFPFELTDNINNHHVFWQRMGR "

ORIGIN

Alignment Scores:

[illegible]

```
QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
Db 181 GCCGCGCGCGCCAGAGAGCTCCACGCGCCAGGAGAACTGGCTGCTCAACAGCAACATC 240
QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
Db 241 CCCCTGGAGACAGAAACCTAGGCAAGCAGCACTTCTAGGACATTCGAGGACAACTC 300
QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGGACATCTCTCGCGCGCTCCGCCCATGATGGCGCTTACAGTACGCGGCACTC 360
QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 361 CTCCTCAAGCTGGAGGACCTCCCGCTCTCTTCGCACCATCACCGGGTACCCACTG 420
QY 306 ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 ACGTCCAGCACAGTGACTCTCTCCGCGCCGACCCCTGCCCGCAGACCTTCGCCCGG 480
QY 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer 345
Db 481 CCGGCTTTTAACTCAAGAGCCCTCCAAAGTACTGTAACCTGGAAGTGGCGACCTGAGC 540
QY 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365
Db 541 GCCATGTCATCTCAGCACCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTG 600
QY 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
Db 601 TTTGGCTAAACTGGACCTGCAGCGATGAGGAGGAGAGATGATGATGATGATGATGATGATG 660
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCGAGCAGTGGCGCTGTGCCAACCGACGCTCTCCCTATACCC 705
```

```
RESULT 13
AB025412 8964 bp mRNA linear ROD 08-MAY-1999
LOCUS Mus musculus mRNA for Ten-m3, complete cds.
DEFINITION AB025412
VERSION AB025412.1 GI:4760779
KEYWORDS Ten-m3.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Ohashi, T., Zhou, X., Feng, K., Richter, B., Moergelin, M., Perez, M.T.,
Su, W., Chiquet-Ehrismann, R., Rauch, U. and Faessler, R.
Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
J. Cell Biol. (1999) In press
2 (bases 1 to 8964)
Ohashi, T.
Direct Submission
Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
(E-mail: ohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128,
Fax: +81-86-222-7768)
FEATURES
Location/Qualifiers
1..8964
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1..8964
/gene="ten-m3"
50..8197
```

gene

CDS

```
/gene="ten-m3"
/codon_start=1
/product="Ten-m3"
/protein_id="BAA77398.1"
/db_xref="GI:4760780"
/translation="MDYKERRPYCSLTKSRREKERRYNTSSADNEECRVPQTKSVSS
ETLKHDIYSRLYGNRKDLVHREADEYTRQGNFTLRQLGVCESTRARGVAFCAE
MGLPHRGYSISAGSDADTENAAMSPEHAMLWGRGVKSRSCLSSRSNLATDTDT
EHNRSDESQSPNNFQPTQLPLPPSHKHQHPHSITSLNRNNTNRNQSPAP
PAALPAELQTPSPESVOLQSWLGNVPLESRHFLFTGTGTTPLFSTATPYTWASG
SVSPPTPLPRLNTLSRAFKKSKSVKRCALCAVGVLLAILLSFIAHLF
GLNHLQQTENDTFENGKNSDTPVNTVPSLPSGDNKLGFTNTHNTHSGELDGR
RAIHEVQPFWRSLFDIQFLKFNLSLQDALIGVYGRKGLPPSHQTDYDFVELLD
GSRILAREQRLNVESEARAGSVSLHEAGFIQYLDLSGIWHLAFYNDKRNPEQSFN
TIVLESVECPNCHNGECVSGTCHGPGFLGPDCAACVLCGNGQYSGKRCCLC
FSWKGTECDVPTTCIDPOCGRGICIMGSCACNKGKGENCEADCLDPCGNCNGV
CIHCECHNCGWGSNCEILKTMCADOSCNGHTYLBQSGSCTCPNWTGPDGCSNEICS
VDCSGHVGWGSRCREHGTGPACORACHPRCAEHGTCCKGKCECSQWNGBHECTI
AHYLDKIVKEGCPGLCNSGRTLDQNGHCVCPQWRGAGCDVAMETLCTGSKDNEG
DGLDMDPDCLOSCCQNPYCRGLPDQDIIQSLOTPSQAQKPYDRISFLIGS
DSTHLPGESPFNKSLASVIRGQVLTADGTPLIGVNVSLFHYSGYVITRQDGMFDL
VANGASLITLIFERSPFLTOVHTVMIWNVYVMDTLVMKEENDISCDLSGFRVS
PLIVSSPLSTFERSSPEDSLIPETQVLHEETIPGTDLSLSLSRAAGTKSVLKIT
MTQAVIPNLKMLHLMVAVNGLRQKMPASPLAYTIFMDKTDAYOKVGLSEAVV
SVGEYESCDLTLWKRKTAVLQGYELDASNGMGTLDKHHVLDVQNGILYKNGENQ
FISQQPVSSIMNGRRSISCPSCQADGNKLAPALACGIDGSLYVYKGNVVR
RIPPSGNVTSVLELRNDRFRHSSPAHYVLTADVTGDLVSDTNRRIYVRPKSLTG
AKDLTKNAEVAGTGEOLPFDEARCGDGKAVBATLMSPKGMALDKNGLIYFVDGTM
IRKVDQNGIISTLLGSNDLTSARPLTCDTSMHISOVLEWPTDLAINPMONSIYVDN
NVVLQITENQVRIRAGRPMHCQVGYEYVPGKHAVTTLTSSATAIYASGVLYITE
TDEKINIRQVITDGEISLVAGIPSECDKNDANDCDYQSGDGYAKDANLAPSSLA
ASPDGTLIADLGNIRAVSKNPLNSMNFVSPVETPDQELYIFDINGHTVTSI
VTDGLYFNSYNDNDVAVTDSNGTLRIRRDNPMRVRVSPDNQVILITGTNGC
LKSMTAQGLBLVFTHGNSGLLATKSDGTFTFDYDGEGLTNVTEFGVYTNLH
GDMKAJTVDISSREEDYSITNSLSDISFYTMQODLNSYQIGYDGLRIPYAS
GLDSHTQTEPHVLACTANPVAKNRMTLPGENQNLVWRFRKEQAGKVNVRGKKUR
VGNRLSLVDFTDKTEKLYDDHRKELLRIAYDTSGHPTLWLPSSKLMANVTYSST
GOIASQRTGTSKVDSDSQSRVSRVADGKTSYTYLTKSVLLHSRQYIFEYD
MMRLSAITMPSVARHMTQIRSIYGYNYNPPESNASITDYNEGILLLOATPLGT
SRVLFKYRQTRLSEILYDSTRVSYTYDETAGVKTNLQSDGFICTIRYOLGLPI
DROIREFSDGMVNARDYSYDMSFRVTSQGVINETPLDILYQFPDDISGVQVQFK
FGVIYDINGLITAVWYTKFPAHGRIRKEIQEYFRSLMYWITITQDNMVRKKE
IKJGPANTTKYAEIVDQQLQTVLNEKIMRYDNLNGHLLNPSSARLTPLR
YLDRLDRIQLGVOYRDEDFLQRGTEIFEYSSKGLLTVYSKSGWTYVRYDGL
GREVSKTSLGQHLQFYADLTPTIRTHVNHSSSTISLYLDLQHLFAMELISGD
EFTYASNTQTPLAFFSSNGLMKQYAYGEIYFDSNVDFOLVIFGHGLYDPLTK
LIHFGDRDYDILAGRWTTPDITWKRIQKDPAPNLYMERNNPNKPIHGVKQYITDV
NSMLVTFGPHLNAIPGFVPKFDLTPEPSVELKSCQWEDYPPIFGVQOQVAKAF
LSLGMKAEVQVSRKAGAESQSLWFAIVKSLIGKVMIAVSQGRVQTNVNIANEDCI
KVAALNNAPYLENLHPTIEGKTHYFIKTTTPESDGLTSLTSGRKALENGINTVS
QSTTVVNGRTRFADVEMQFALALHVRVYGMTLDEEKARILEQARALARAWAREQ
RVRDGEGARLWTEGEKQLLSAGKVGQYGVYVLSVEQYPELADSANNTQFLRQSEI
GKR"
```

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 1.67e-42 | Length: | 8964 |
| Score: | 982.00 | Matches: | 210 |
| Percent Similarity: | 62.65% | Conservative: | 50 |
| Best Local Similarity: | 50.60% | Mismatches: | 91 |
| Query Match: | 46.06% | Indels: | 64 |
| DB: | 10 | Gaps: | 12 |

SEQ14-X-AT-28-64-76 (1-400) x AB025412 (1-8964)

| | | | |
|----|-----|--|-----|
| QY | 1 | MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu | 19 |
| Db | 50 | AUGCAITGTAAGGACGACGAGCTTACTGCTCTTACCAAGAGCAGACGGGAAGGAA | 109 |
| QY | 20 | ArgArgTyrThrSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer | 38 |
| Db | 110 | AGGCGCTATACAAATTCGTCGCGGCAANTGAGAGTGTAGGGTCCCCACGACAGAGTCC | 169 |

```
QY 39 TyrSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 170 TATAGTTCAGTGAAGAACTTGAAGACTTCGATCATGATTATTCACGGCTGCTTATGGA 229
QY 58 SerArgValLysAspIle***ProGlnGluAlaGluPheCysArgThrGly***Asn 77
Db 230 AACAGAGTAAAGATTGGTCCACAGAGAAGCGGAGTAGTATACAGACAGACAGAAAT 289
QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArg 97
Db 290 TTTACCTTAAGCAGTGTAGGAGTGTGTGAATCCGCAACTCGAAGAGGAGTGCATCTCT 349
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 350 GCGGAATGGGCTCCCTCACAGAGTTACTCATCAGTGCAGGTCAGATCGGATACG 409
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 410 GAAACGAGACAGTGTATCCCTGAGCATGCGCATGAGACTTTGGGGCAGGGGGTCAAA 469
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 470 TCGGGCGGAGTTCCTCGCTGTCAAGCGGTCAACTCCGCCCTCACCTGCAGACAG 529
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 530 GAGCAGGAGAAGTTCGAGCAGTGAAGGAGGAGCACTTCAACACCCAGGGCAACCC 589
QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisAla 194
Db 590 ACCCTGAGCTTTGCGGCA-----TCCACAGCAGCACCCTGCGGAGCATCAC--- 640
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 641 CGCTCCATCACTTCCTCAATAGAACTCCCTGACCAATAGAGGAAACAGATCCGGCC 700
QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 701 CCG-----CCGGCTGCTTTGCGCGGAGTGTCAACACCAACCA 736
QY 231 ---GluProAlaHisAlaGlnGluAsnThrLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 737 CCGCAGTCCGTCAGCAGTGCAGGACAGTGGTGTCTTGGCAGTAAATGACACTGGAAGC 796
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 797 AGG----- 799
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 800 -----CATTTCTTATCAAAACA 817
QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 818 GGGACAGGAGCAGCAGCTGTTTTCAGTACGCGCAACCCCGGAGTACACAATGGCATCTGGC 877
QY 309 ThrValTyrSerProProProProProLeuProLeuProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 878 TCTGTTTATTTCTCCGCTACCGGCACTTCTCTAGAAACCCCTATCAAGAGTGCTTTT 937
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 938 AAATTCAGAGTCTTCAAGTAGTACTGAGCTGGAGGTGCACCGCACTGTGTGCTGTAGGG 997
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 998 GTCTCAGTGTCTGGCCATTTCTCTCTATTTATAGCATGCAATCTATTGTCCTC 1057
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1058 AACTGGCATTACAGCAGCGGAAATGACACATTCGAGATCGAAAGTGAATCTTGAC 1117
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
```

```
Db 1118 ACC-----GTGCCAAACAACTGATGATGCTTACCT 1147
RESULT 14
AF195420 1476 bp mRNA linear PRI 27-JAN-2000
LOCUS Homo sapiens OD23 (OD23) mRNA, partial cds.
DEFINITION AF195420
ACCESSION AF195420
VERSION AF195420.1 GI:6760372
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1476)
AUTHORS Ben-Zur, T., Feige, E., Motro, B. and Wides, R.
TITLE The mammalian Odz gene family: Homologs of a Drosophila pair rule gene with expression implying distinct yet overlapping developmental roles
JOURNAL Dev. Biol. 117, 107-120 (2000)
REFERENCE 2 (bases 1 to 1476)
AUTHORS Ben-Zur, T., Motro, B. and Wides, R.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-1999) Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel
FEATURES
Location/Qualifiers
1..1476
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>1476
/gene="OD23"
/notes="Ten-m3"
151..>1476
/gene="OD23"
/notes="similar to the Drosophila melanogaster pair rule odd Oz/Ten-m gene sequence"
/codon_start=1
/product="OD23"
/protein_id="AAF28318.1"
/db_xref="GI:6760373"
/translation="MDVKRRPYCSLTKSREKERYTNSADNECRVPTOKSYSSS
ETLAFDHSRLLYGNRVKDLVHREADFTQGNFTLRQLGVCSPATPRGLAFCAE
MGLPHRGYSISAGSDADTNEAVSPHAWRLMGRGVKSGRSCLSRSNSALTLDIT
EHNKSDSNEQPSAQGSTLOPLPEKSHQSAHPHSTITSLNRLNINRRNQSPAP
PALPAELQTTPEVOLQSWLGSNVPLESRHFLFKTGTGTTPLSTATPGYTMASG
SVYSPTRPLPRNTLSRSAPFKKSKYCSWKCTALCAVGSVLLAILLSYFIAMHLF
GLNWLOQTENDTFENGKYNSTMTNTVSLPSGDNKGLGFTQENNTIDSGELDIGR
RAIQEIPPGIFWRSQLFIDQOPQLKFNISLQKDALIGVYGRKKLPESHQ"
ORIGIN
Alignment Scores: 5.5e-43 Length: 1476
Pred. No.: 975.00 Matches: 210
Score: 62.41% Conservative: 49
Percent Similarity: 62.41% Mismatches: 92
Best Local Similarity: 50.60% Indels: 64
Query Match: 45.73% Gaps: 12
DB: 9
SEQ14-X-AT-28-64-76 (1-400) x AF195420 (1-1476)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGGATGTGAAGAAACGACGCGCTTACTGCTCCTGACCAACATGAGGAGTCCGGGTACCCACAGAGTCC 210
QY 20 ArgArgTyrThrSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGGCGCTACACAAATTCCTCCGACAGCAATGAGGAGTCCGGGTACCCACAGAGTCC 270
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTTCACGAGACATTTGAAAGCTTTTTCATCATGATTCCTCGCGGCTGCTTACGGC 330
```

```

QY 58 SerArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
Db 331 AACAGAGTGAAGATTGGTTACAGAGAGACGACGAGTTCTAGACAGAGCAGAAAT 390
QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
Db 391 TTACCCCTAAGCAGTTAGGAGTTGTGAACAGCAACTCGAAGAGAGTGGCATTTTGT 450
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGGAATGGGGCTCCCTCACAGAGTTACTATCATGTCGAGGTCAGATGCTGATACT 510
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 511 GAAATGAAGCAGTATGCTCCACAGAGCATGCCATGACACTTTGGGGCAGGGGTCAA 570
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCGCAGCTCTGCTGTCAAGTCGGTCCAACTCAGCCCTCACCTCGACAGATACG 630
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAACAAAGTCGACAGTGAAGTGAACCTGCAAGCAATCAAGGCGCAGTCT 690
QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisAla 194
Db 691 ACCCTGCAGCCCTTGGCGCT-----TCCATTAAGCAGCACTCTGCACAGCATCAT 741
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCATCACTTCTCTCAACAGAAACTCCCTGACCAATAGAGGAACCCAGAGTCGGCC 801
QY 215 ProThrAspHisSerLeuSerGlyLeuProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CCGCTCTTTGGCCGCGAGCTGCAACACCA 837
QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCGAGTCCGTCAGCTGCAGGACAGTGGTCTTGGCGATATGATACCACTGGAAGC 897
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 898 AGG----- 900
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTTCTTATTCAAACA 918
QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGGTACAACGCCACTGTTTCAGTACTGCAACCCAGAGTACAAATGGCATCTGGC 978
QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTACACCACTACTCGGCCACTACTAGAAACACCCCTATCAAGAGTGCTTTT 1038
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAGAAGTCTTCAAGTACTGTAGCTGAAATGCACTGCTGTGCGGTAGGG 1098
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMethHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCTCGCAATACCTCTGCTTATTTATAGCAATGATCTCTTTGGCCCTC 1158
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGCGACCTACAGCAGACTGAAATGACATTTGAGAATGAAAGTGAATTCGTAT 1218
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1219 ACC-----ATGCCAACAAACACTGTGTCTATTACCT 1248

```

RESULT 15

```

AX662355
LOCUS AX662355 8645 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 37 from Patent WO02062999.
ACCESSION AX662355
VERSION AX662355.1 GI:29163217
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
Malayankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,
Gorman, L., Kkuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
Millett, I. and Macdougall, J.R.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 02062999-A 37 15-AUG-2002;
FEATURES
source
1. 8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

Alignment Scores:

```

Pred. No.: 3,76e-42 Length: 8645
Score: 975.00 Matches: 210
Percent Similarity: 62.41% Conservative: 49
Best Local Similarity: 50.60% Mismatches: 92
Query Match: 45.73% Indels: 64
DB: 6 Gaps: 12

```

SEQ14-X-AT-28-64-76 (1-400) x AX662355 (1-8645)

```

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgAspAlaGlu 19
Db 151 ATGGATGTGAAGAACGACGAGCCCTTACTGCTCCTGACCAAGAGACGACGAGAGGAA 210
QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGCGCTTACAAATTCCTCCGACAAATGAGAGGTGCGGGTACCCACACAGAGTCC 270
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTTCACGACGACATTTGAAGCTTTTGATCATGATTCCTCGCGCTGCTTTACGC 330
QY 58 SerArgValLysAspIle***ProGlnGluAlaGluPheCysArgThrGly***Asn 77
Db 331 AACAGAGTGAAGATTGGTTACAGAGAGACGACGAGTTCTAGACAGAGCAGAAAT 390
QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArg 97
Db 391 TTACCCCTAAGCAGTTAGGAGTTGTGAACCAACGACCACTCGAAGAGAGTGGCATTTGT 450
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyValaGlySerAspAlaAspMet 117
Db 451 GCGGAATGGGGCTCCCTCACAGAGTTACTATCATGTCGAGGTCAGATGCTGATACT 510
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 511 GAAATGAAGCAGTATGCTCCACAGAGCATGCCATGACACTTTGGGGCAGGGGTCAA 570
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCGCAGCTCTGCTGTCAAGTCGGTCCAACTCAGCCCTCACCTCGACAGATACG 630
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAACAAAGTCCGACAGTGAAGTGAACCTGCAAGCAATCAAGGCGCAGTCT 690

```

```

QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db 691 ACCCTGCAGCCCTTGCOCCTT-----TCCCATAGCAGCAGCTCTGCACAGCATCAT--- 741
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCATCACTTCTCTCAACAGAACTCCCTGACCAATAGAGGAAACCAGAGTCCGGCC 801
QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CCGGCTGCTTTGCCCGCCGAGCTGCAAAACCACA 837
QY 231 ---GluProAlaHisAlaGlncluhAsnTrpLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCCGAGTCGTCACAGTCGAGCAGCTGGGCTCTGGCAGTAAATGATGACCTGGAAAGC 897
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 898 AGG----- 900
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTTCTATTCAAAACA 918
QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGGTACACCGCCTGTTTCTAGTACTGCAACCCCGAGGATACACATGGCACTGGC 978
QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTCACCACTTCTCGGCCACTACCTAGAAACACCCCTATCAAGAAGTGCTTTT 1038
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAAGAGTCTTTCAAAAGTACTGTAGCTGAAATGCACCTGCTGTGCCGTAGGG 1098
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMethHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCTGCAATATCTCTCTCTTATTTATAGCAATGATCTCTTTGGCCCTC 1158
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGTACAGCAGACTGAAATGACACATTTGAGAAATGAAAAAGTGAATCTGAT 1218
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1219 ACC-----ATGCCAACAAACACTGTGTCAATACCT 1248

```

Search completed: August 14, 2004, 23:48:40
Job time : 5209 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 1255.76 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_450_520

Perfect score: 372
Sequence: 1 TFWRSQVFIDHPVHLKFNVS.....RLLTQEARSLGTPRQSRGT 71

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-OUTFMT=ptc -NORW=ext -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-USER=US10029020@cgn 1 1 19065 @runat 06082004 112216 29275 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOJURY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_btg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|-------|---------------|--------|----|-----------|---------------------|
| 1 | 372 | 100.0 | 1680 | 6 | AR270935 | AR270935 Sequence |
| 2 | 372 | 100.0 | 2387 | 6 | AR270936 | AR270936 Sequence |
| 3 | 372 | 100.0 | 3111 | 6 | BD270887 | BD270887 Method of |
| 4 | 372 | 100.0 | 3111 | 6 | AR270934 | AR270934 Sequence |
| 5 | 372 | 100.0 | 3111 | 6 | BD140153 | BD140153 Utilizati |
| 6 | 372 | 100.0 | 3111 | 9 | AF009227 | AF009227 Homo sapi |
| 7 | 372 | 100.0 | 8354 | 6 | AX556500 | AX556500 Sequence |
| 8 | 372 | 100.0 | 8438 | 6 | AX675551 | AX675551 Sequence |
| 9 | 372 | 100.0 | 8645 | 6 | AX600210 | AX600210 Sequence |
| 10 | 353 | 94.9 | 8585 | 10 | AB025413 | AB025413 Mus muscu |
| 11 | 347 | 93.3 | 9722 | 10 | AF059485 | AF059485 Mus muscu |
| 12 | 277 | 74.5 | 9264 | 5 | AB026980 | AB026980 Dario rer |
| 13 | 254 | 68.3 | 8473 | 6 | AX662357 | AX662357 Sequence |
| 14 | 254 | 68.3 | 8487 | 6 | AX662359 | AX662359 Sequence |
| 15 | 254 | 68.3 | 8645 | 6 | AX662355 | AX662355 Sequence |
| 16 | 248.5 | 66.8 | 7816 | 10 | AF195418 | AF195418 Mus muscu |
| 17 | 248.5 | 66.8 | 8816 | 5 | AB026979 | AB026979 Dario rer |
| 18 | 248.5 | 66.8 | 8964 | 10 | AB025412 | AB025412 Mus muscu |
| 19 | 234 | 62.9 | 8675 | 6 | AX662353 | AX662353 Sequence |
| 20 | 232.5 | 62.5 | 17356 | 2 | AC134948 | AC134948 Rattus no |
| 21 | 231.5 | 62.2 | 93474 | 2 | AC136801 | AC136801 Rattus no |
| 22 | 231.5 | 62.2 | 245722 | 2 | AC120288 | AC120288 Rattus no |
| 23 | 228 | 61.3 | 226256 | 10 | AC073599 | AC073599 Mus muscu |
| 24 | 222.5 | 59.8 | 8118 | 5 | GGA238613 | GGA238613 Gallus ga |
| 25 | 222 | 59.7 | 165812 | 2 | AP001141 | AP001141 Homo sapi |
| 26 | 222 | 59.7 | 166063 | 9 | AP002515 | AP002515 Homo sapi |
| 27 | 222 | 59.7 | 183308 | 2 | AC084775 | AC084775 Homo sapi |
| 28 | 222 | 59.7 | 175464 | 2 | AP002887 | AP002887 Homo sapi |
| 29 | 220.5 | 59.3 | 8373 | 10 | AB025410 | AB025410 Mus muscu |
| 30 | 220 | 59.1 | 8297 | 9 | AF100772 | AF100772 Homo sapi |
| 31 | 194.5 | 52.3 | 1430 | 6 | AX250006 | AX250006 Sequence |
| 32 | 194.5 | 52.3 | 1431 | 6 | AX058614 | AX058614 Sequence |
| 33 | 194.5 | 52.3 | 1727 | 6 | AX714026 | AX714026 Sequence |
| 34 | 194.5 | 52.3 | 1727 | 9 | AX056053 | AX056053 Homo sapi |
| 35 | 194.5 | 52.3 | 1743 | 6 | AX133821 | AX133821 Sequence |
| 36 | 194.5 | 52.3 | 2496 | 5 | GGA245711 | GGA245711 Gallus ga |
| 37 | 194.5 | 52.3 | 2496 | 6 | AX250064 | AX250064 Sequence |
| 38 | 194.5 | 52.3 | 7713 | 9 | HSM808325 | HSM808325 Homo sapi |
| 39 | 194.5 | 52.3 | 8409 | 5 | GGA279031 | GGA279031 Gallus ga |
| 40 | 194.5 | 52.3 | 8409 | 6 | AX250068 | AX250068 Sequence |
| 41 | 194.5 | 52.3 | 8575 | 6 | AX921803 | AX921803 Sequence |
| 42 | 194.5 | 52.3 | 8689 | 6 | AX250067 | AX250067 Sequence |
| 43 | 194.5 | 52.3 | 8689 | 10 | AF086607 | AF086607 Rattus no |
| 44 | 194.5 | 52.3 | 8797 | 6 | AX250063 | AX250063 Sequence |
| 45 | 194.5 | 52.3 | 8797 | 6 | AX250066 | AX250066 Sequence |

ALIGNMENTS

RESULT 1

```
AR270935
LOCUS AR270935 1680 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 3 from patent US 6500941.
ACCESSION AR270935
VERSION AR270935.1 GI:29702177
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 1680)
AUTHORS Schaefer,G.M. and Sliwkowski,M.
TITLE Gamma-heregulin
JOURNAL Patent: US 6500941-A 3 31-DEC-2002;
FEATURES
  Location/Qualifiers
    1..1680
      /organism="unknown"
      /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 6.57e-35 Length: 1680
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x AR270935 (1-1680)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
DB 1348 ACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCATCTGAAATTCATGTGCT 1407
QY 21 LeuGlyLysAlaAlaValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
DB 1408 CTGGGAAAGGCGCCCTGTTGGCATTATGGCAGAAAAGGCTCCCTCTTCACATACA 1467
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
DB 1468 CAGTTTGACTTTGTGAGCTGCTGGATGGCAGGAGCTCCTAACCCAGGAGCGGGAGC 1527
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
DB 1528 CTAGAGGGAGCCCGCGCCAGTCTCGGGGAAGT 1560
RESULT 2
AR270936
LOCUS AR270936 2387 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 11 from patent US 6500941.
ACCESSION AR270936
VERSION AR270936.1 GI:29702178
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 2387)
AUTHORS Schaefer,G.M. and Sliwkowski,M.
TITLE Gamma-heregulin
JOURNAL Patent: US 6500941-A 11 31-DEC-2002;
FEATURES
  Location/Qualifiers
    1..2387
      /organism="unknown"
      /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 9.32e-35 Length: 2387
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
```

```
US-10-029-020-14_COPY_450_520 (1-71) x AR270936 (1-2387)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
DB 853 ACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCATCTGAAATTCATGTGCT 912
QY 21 LeuGlyLysAlaAlaValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
DB 913 CTGGGAAAGGCGCCCTGTTGGCATTATGGCAGAAAAGGCTCCCTCTTCACATACA 972
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
DB 973 CAGTTTGACTTTGTGAGCTGCTGGATGGCAGGAGCTCCTAACCCAGGAGCGGGAGC 1032
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
DB 1033 CTAGAGGGAGCCCGCGCCAGTCTCGGGGAAGT 1065
RESULT 3
BD270887
LOCUS BD270887 3111 bp DNA linear PAT 17-JUL-2003
DEFINITION Method of promoting the proliferation of inner ear hairy cells
  using ligand of HER2 receptor and/or HER3 receptor.
ACCESSION BD270887
VERSION BD270887.1 GI:33080655
KEYWORDS JP 2002529425-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 3111)
AUTHORS Gao,W.O.
TITLE Method of promoting the proliferation of inner ear hairy cells
  using ligand of HER2 receptor and/or HER3 receptor
JOURNAL Patent: JP 2002529425-A 6 10-SEP-2002;
COMMENT
  GENE-TECH INC
  OS Homo sapiens (human)
  PN JP 2002529425-A/6
  PD 10-SEP-2002
  PF 28-OCT-1999 JP 2000580655
  PR 07-NOV-1998 US 60/107522
  PI WEI QIANG GAO
  PC A61K45/00,A61K38/00,A61K39/395,A61P27/16,C12N5/06,C12N5/06,
  C12R1/91,
  PC A61K37/02,C12N5/00,(C12N5/00,C12R1/91)
  CC Method of promoting the proliferation of
  inner ear hairy cells
  CC using
  CC ligand of HER2 receptor and/or HER3 receptor
  FH Key Location/Qualifiers
  FT source 1..3111
  FT /organism='Homo sapiens (human)'
  FT Location/Qualifiers
  1..3111
    /organism='Homo sapiens'
    /mol_type='genomic DNA'
    /db_xref='taxon:9606'
ORIGIN
Alignment Scores:
Pred. No.: 1.21e-34 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x BD270887 (1-3111)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
DB 1681 ACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCATCTGAAATTCATGTGCT 1740
```


QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
 Db 1741 CTGGGAAGGAGCGCTGGTTGGATTATGGCAGAAAAGCCCTCCCTTCATACATA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
 Db 1801 CAGTTTGACATTGTGGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
 Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1893

RESULT 4
 LOCUS AR270934 3111 bp DNA linear PAT 10-APR-2003
 DEFINITION Sequence 1 from patent US 6500941.
 ACCESSION AR270934
 VERSION AR270934.1 GI:29702176
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 3111)
 AUTHORS Schaefer,G.M. and Sliwowski,M.
 TITLE Gamma-herregulin
 JOURNAL Patent: US 6500941-A 1 31-DEC-2002;
 FEATURES Location/Qualifiers
 source 1..3111
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.21e-34 Length: 3111
 Score: 372.00 Matches: 71
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AR270934 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
 Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTGATCTGAAATTCATGTCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
 Db 1741 CTGGGAAGGAGCGCTGGTTGGCATTTATGGCAGAAAAGCCCTCCCTTCATACATA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
 Db 1801 CAGTTTGACATTGTGGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
 Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1893

RESULT 5
 LOCUS BD140153 3111 bp DNA linear PAT 18-SEP-2002
 DEFINITION Utilization of heregulin as epithelial cell growth factor.
 ACCESSION BD140153
 VERSION BD140153.1 GI:23235098
 KEYWORDS JP 2002509076-A/6.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 3111)
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 TITLE Sliwowski,M. and Kern,J.A.
 JOURNAL Utilization of heregulin as epithelial cell growth factor
 Patent: JP 2002509076-A 6 26-MAR-2002;

COMMENT
 OS Homo sapiens (human)
 PN JP 2002509076-A/6
 PD 26-MAR-2002
 PF 03-FEB-1999 JP 2000530226
 PR 04-FEB-1998 US 09/020598
 PI MARK Sliwowski,JEFFREY A KERN
 PC A61K38/00,A61K35/12,A61K39/395,A61K39/395,A61P11/00,A61P43/00,
 PC C07K14/485,
 PC C07K16/22,C12N15/09,C12P21/08,A61K37/02,C12N15/00 CC
 Utilization of heregulin as epithelial cell growth factor PH key

FT source 1..3111
 FT Location/Qualifiers
 /organism="Homo sapiens (human)"

FEATURES
 source 1..3111
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.21e-34 Length: 3111
 Score: 372.00 Matches: 71
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x BD140153 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
 Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTGATCTGAAATTCATGTCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
 Db 1741 CTGGGAAGGAGCGCTGGTTGGCATTTATGGCAGAAAAGCCCTCCCTTCATACATA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
 Db 1801 CAGTTTGACATTGTGGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
 Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1893

RESULT 6
 LOCUS AF009227 3111 bp mRNA linear PRI 17-SEP-1997
 DEFINITION Homo sapiens gamma-herregulin mRNA, complete cds.
 ACCESSION AF009227
 VERSION AF009227.1 GI:2406643
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 3111)
 AUTHORS Schaefer,G., Fitzpatrick,V.D. and Sliwowski,M.X.
 TITLE Gamma-herregulin: a novel heregulin isoform that is an autocrine growth factor for the human breast cancer cell line, MDA-MB-175
 JOURNAL Oncogene 15 (12), 1385-1394 (1997)
 MEDLINE 97472144
 PUBMED 933014
 REFERENCE 2 (bases 1 to 3111)
 AUTHORS Schaefer,G., Fitzpatrick,V.D. and Sliwowski,M.X.
 TITLE Direct Submission
 JOURNAL Submitted (18-JUN-1997) Protein Chemistry, Genentech, Inc., 460 Point San Bruno Blvd, South San Francisco, CA 94080, USA

FEATURES
 source 1..3111
 Location/Qualifiers
 /organism="Homo sapiens"

```
/mol_type="mrna"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22-p11"
/cell_line="MDA-MB-175"
/cell_type="breast cancer cell"
334.._2640
/codon_start=1
/product="gamma-heretulin"
/protein_id="AAC51756.1"
/db_xref="GI:2406644"
/translation="MDVKRPYSLRRRDAERYTSSADSEBGAPOKSYSSST
LKAYDQDLARVGRVQIVQEAEEFCRTGANFTLRELGLLEVPVPHGLTYRTDIGL
PHCYSGAGSDADMEATVLSPEHVLWGRSTESGRSSCLSSPANSNLTLTDTHE
NRETDHPGQNHARLRTPPPLSHAHVPHQHAASINSNLRGNETPRSNRSPATDTH
SLSGPPAGGAQEPHAHGNLLNSNIPLETRNLGKQPLGLTLQNLLEMDILGASRH
KCYWGDHFLPKGTSPLCTFCSTYSPSTVSPPLPRSTPAAPNPKKPS
DGNWCAALSAIVISATLIVLAVFAMHLFGLNHLQPMGQMYEITEDTASWPV
PTDVSYPGGTLETPDKRGKTTGKSPSPFPEDSFIDSGEIDVGRRASOKIPGT
PWRQVFTDHPVHLKFNVSLGKAALVGVGRKLPSPHTQDFVELLGRLLTCEAR
SLEGTPRSGRTVPSSSHETGFIQVLDGSIWHLAINDKSESVFLLTALALPRL
KEMQSQAASGKVLRCETSEYSSSLRPFKFNKLNELNKNKPNKIQRKPGKSEL
RINKASLADSGEYMKCVISKLNDGSANITIVESNEITGMPTAGTEGAYVSESPIR
ISVSTEGANTSSSTSTGTGSHLVKCAEKEKTCVNGGCECFMWKDLNPNRSLCKCP
NEFTGDRQCNVMAFSYSTSTPFLSLPE"
ORIGIN
Alignment Scores:
Pred. No.: 1.21e-34 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x AF009227 (1-3111)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTCATAGACCATCTGTCATCTGAAATTCATGTGCT 1740
QY 21 LeuGlyLysAlaLeuValGlyLysGlyArgGlyLeuProProSerHisThr 40
Db 1741 CTGGAAAGCGACCCCTGTTGGCATTTATGCGAGAAAGGCTCCTCCTTCACATACA 1800
QY 41 GlnPheAspPheValGlnLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGAGCTGCTGATGCGAGAGGCTCCTAACCCAGGAGCGCGGAGC 1860
QY 61 LeuGlyLysThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGAGCCCGCGCCAGCTCTCGGGGAAT 1893
RESULT 7
AX556500
LOCUS AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO2057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyankar,U.M.,
Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E.,
Zerhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.I.,
Smithson,G., Li,L. and Ji,W.
Polypeptides and nucleic acids encoding same
Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
Location/Qualifiers
FEATURES
ORIGIN
Alignment Scores:
Pred. No.: 3.27e-34 Length: 8438
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x AX675551 (1-8438)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1382 ACTTCTGGAGATCTCAAGTGTCATAGACCATCTGTCATCTGAAATTCATGTGCT 1441
QY 21 LeuGlyLysAlaLeuValGlyLysGlyArgGlyLeuProProSerHisThr 40
Db 1442 CTGGAAAGCGACCCCTGTTGGCATTTATGCGAGAAAGGCTCCTCCTTCACATACA 1501
QY 41 GlnPheAspPheValGlnLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 1502 CAGTTTGACTTTGTGAGCTGCTGATGCGAGAGGCTCCTAACCCAGGAGCGCGGAGC 1561
QY 61 LeuGlyLysThrProArgGlnSerArgGlyThr 71
Db 1562 CTAGAGGGAGCCCGCGCCAGCTCTCGGGGAAT 1594
RESULT 8
AX675551
LOCUS AX675551 8438 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 1 from Patent WO2055704.
ACCESSION AX675551
VERSION AX675551.1 GI:29333552
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Padigaru,M., Li,L., Zerhusen,B.D., Casman,S.J., Shenoy,S.,
Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M.,
Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E.,
Boldog,F.I., Gross,W.M., Alsobrook,J.P., Gerlach,V.,
Edingermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J.,
Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and
Stone,D.J.
Proteins, polynucleotides encoding them and methods of using the
same
Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)
Location/Qualifiers
FEATURES
ORIGIN
Alignment Scores:
Pred. No.: 3.27e-34 Length: 8438
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x AX675551 (1-8438)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
```

```
Db 1358 ACTTCTCGAGATCTCAAGTTCATAGACCACTCTGTGCACTCGAATTCATATGCTCT 1417
QY 21 LeuGlyLeuAlaalaLeuValGlyLeuTyrGlyArgLeuGlyLeuProSerHisThr 40
Db 1418 CTGGGAAAGCAGCCCTGTTGGCAATTTATGGCAGAAAGCCCTCCTCTCCACATACA 1477
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgLeuLeuThrGlnGluAlaArgSer 60
Db 1478 CAGTTTGACTTTGTGAGCTGCTGGATGGCAGAGCTCTTAACCCAGGAGGCGCGAGC 1537
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1538 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1570

RESULT 9
AX600210 8645 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S.,
Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,
Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K.,
Khare,R. and Walla,N.K.
TITLE Proteins associated with cell growth, differentiation, and death
JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
INCYTE Genomics, Inc. (US)
FEATURES
source
1. .8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/notes="Incyte ID No: 7488573CB1"

ORIGIN
Alignment Scores:
Pred. No.: 3,35e-34 Length: 8645
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AX600210 (1-8645)
QY 1 ThrPheTyrArgSerGlnValPheIleAspHisProValHisLeuIysPheAsnValSer 20
Db 1452 ACTTCTCGAGATCTCAAGTTCATAGACCACTCTGTGCACTCGAATTCATATGCTCT 1511
QY 21 LeuGlyLeuAlaalaLeuValGlyLeuTyrGlyArgLeuGlyLeuProSerHisThr 40
Db 1512 CTGGGAAAGCAGCCCTGTTGGCAATTTATGGCAGAAAGCCCTCCTCTCCACATACA 1571
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgLeuLeuThrGlnGluAlaArgSer 60
Db 1572 CAGTTTGACTTTGTGAGCTGCTGGATGGCAGAGGCTCTTAACCCAGGAGGCGCGAGC 1631
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1632 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1664

RESULT 10
AB025413 8585 bp mRNA linear ROD 08-MAY-1999
LOCUS Mus musculus mRNA for Ten-m4, complete cds.
DEFINITION
ACCESSION AB025413
```

```
VERSION AB025413.1 GI:4760781
KEYWORDS Ten-m4.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Ohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T.,
Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
TITLE Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
JOURNAL J. Cell Biol. (1999) In press
AUTHORS J. (bases 1 to 8585)
Ohashi,T.
DIRECT SUBMISSION
SUBMITTED (29-MAR-1999) Toshitaka Ohashi, Okayama University
MEDICAL SCHOOL, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
(E-mail: ohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128,
Fax: +81-86-222-7768)
FEATURES
Location/Qualifiers
1. .8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1. .8585
/gene="ten-m4"
188. .8503
/gene="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"
/translat="MDVKERKPVSLTRRDAERRVTSSSADSECKGPKQKSSSET
LKAYQDARLAYGSRVKMVPQAEFCCTGTFNRLGLGEMTPHGLTRTDGL
PHGYSGLASSADLEADTVLSEPHVRLWGRSTRGRSSCLSSRANSNLITDDEH
NTETHPSLSLQNHRLRTPPLPHATFNQHHASINSLNRGNFTPRNPSPTDTH
SLSGEPAGAQEPHTAQDNWLNLSNIPLETNRLKQPLFTLQDNLIEMLILASRH
DGAYSDFGLFKFGGTSPLFTTSPGYLTSTVSPPLPRFSPAFNLKPS
KYCNKCAALISATIVILLAYFVAMHLGLNWLHLOPMEGOMQVETTEDTASSV
PVYDVLSYSGGTGLETPDRKGAEGKPSLSPEDSFIDSGEIDVGRRAQKPP
GTFNRSQVFLDHPHLKFNVLGKALGLYGRKGLPPSHQTPDFVELLDGRLLQEQ
ARSLGPGQSRGPPVPPSHETGFIQYLDSGIWHLAFYNDGKESVSVSLTAAESVD
NCPNCGYGDICISGTCFLGFLGCRASCPVLCGNGQYMKRCCLCHSGKGAEL
CVPTNQCIDVACSSHGTCIMGTICINPGYKGESCEEVDCMDPTCSRSRGVYRGECH
SVYGGTNCETPRATCLDQSGHGTFLPDTGLNCDDPSWTGHDCSI CAADCGHGV
CVGGTCREDGMMGAACDORACHPRCAEHGTCRDKCECSGWNHGECTIAHYLDREV
KSGPCLNGNRCITLDLNGHVCYCOLGWRGTGCTSMETGGDGKDNDDGLVDGMD
POCCPLCHVNPCLGSPDLDTIOETQAPVSQQNLNPFYDRIKFLVGRDSTHSIPG
ENPFQGHACVIRGQVMTSDGTPLVGNISFINNPLFGYI SRQDGSFLLVNGGFSI
ILRFRAPIITQETHLWLPDRFVMTETVMRHEENIEPSCDLSNFAFPNVPVSPSL
TFASACBKGPIVPEIOALQOEIVTAGKRLSYLSRTPGKSVLRISLTHPTIPF
NLMKVLMTVAVSGRFRKFAAPDLVYFIMDKTDVYNQKVGFSFAVSVGYEVES
CEDLILWEKRTAVLQGYEIDASKLGSLDKHNLNIQSGILHKNGENQFVSQQPVV
IGSIMNGRRRSISCSNGLADGNKLAPVALTCGDSGLYVGDENYIRRIFFPSNV
TNILMRNDRFRHSHPAKYYIATDPMGAVFLSDTNRRRVFKVSTTWKDLVKN
EVVAGTDCDLPDDTCDGGKATEATLNPRTGIVTKFGLIYFVDGTMRVVDQNG
IISTLLGNDLTSARPLSCDSVMEISOVRLEWPTDLAINPMDNSLVLDNNVYLOISE
NHQVRIVAGRPMHCQVPGIDHLLSKVAJHATLESATALAVSHNGVLYATDEKKN
RQVTTSEISILVAGPSGCCDKDANCFSGDDGYAKADKLNLPSSLAVCAOGEL
YVADLGNIRIRIRNNKPFNTNMYELUSSPIDQLYLDFTSGKHLTQSLPTGGLY
NFTYTGDDGDIHTIDNNGMNVNRDSTGMLWLVVDGQVYVMTGNTSALRSVTQ
GHELAMTYHNGSLIATKSNENGWTFYEDYFGLRTNVTPTGVSFSRSDTSV
HVQVETS SKDDVTITTNLSAGAFVTLLOQVRNSVYICADGSLRLLLANGMEVLAQT
EPHLLAGTVNPTVKENVTLPIDNGLNVEMQRKEAQGVQADQVTFEERLVHNNLLS
LODFRVTRTEKIYDHRKFTLRILYDQAGPSLWSPSSRLNGVNVYIFGGLHAGLQR
GIMSMEYDQAGRTSRIFADGKMSVYILEKSMVLHLHSQRYIFEDKNDRIQR
TMPNARQLETIRSVGYRNYQPEGNASVQIDFTDHLHLHTFLGTRRVIKY
GKLSKLAETLYDTTKVSTYDGTAGMLKTNLQNEGTCTIRYRQIGPLIDRQIFPT
ESGMVNAFDYNDNSFRVTSMQAVINETPLPIDLYRYDSDVSGTKQFGKFGVYIDI
```

NQIITTAAMTHKHFDAYGRMKVEQYEIFRSLMYMTVQYDNGRVRVKKELKGVYAN
TRYSEYVADGQLOQTVSINDKPLWRSYDNLGNLHLLSPGNSARLTLRLYDLDRIT
RLGVQVQWMDGQLOQTVSINDKPLWRSYDNLGNLHLLSPGNSARLTLRLYDLDRIT
SHSHLQFYADLNPNTKVTLYNHSSEITSLYDLOCHLFAVELSGDFEYIACDN
IGTLAVFSGTGLMIKQILYAYGIBYMTNPNFQIILGYHGGYDPLTKLVHMGRRD
YDVLAWRTSPDHELMKRLSSNIVPFLYKNNPISNSQDIKCFMTDVSMLLTF
GFQHLNVPKPKPTDAMEPSYELVHTQMTQWDSKSLGVQCEVQKQKAFVTL
ERFQLYGTTTSCQQAPETKKFASSGSIKGVKFKALDKGRVTVDIIISVANEGRRI
AAILNNAHLENLHFTIDGVTHYFVKPSEGLAILGLSGRRTLENGVNTVSQI
NTMUGRRRTYDLOQFALCLNTRYGTVDDEKRVLELARAQVRAQWARQOORL
REGREGUAWTDGKQOVLNTRGVQYDGFVTSVEQYPELSDSANNIHFMRQSEMR
R"

ORIGIN

Alignment Scores:
Pred. No.: 6,28e-32 Length: 8585
Score: 353.00 Matches: 68
Percent Similarity: 97.14% Conservative: 0
Best Local Similarity: 97.14% Mismatches: 2
Query Match: 94.8% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AB025413 (1-8585)

Qy 1 ThrPheTpaArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1541 ACGTTCGAGATCTCAGGTGTCATAGACACCTGTACACCTGATCAATGCT 1600
Qy 21 LeuGlyLysAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 1601 CTGGGGAAGCAGCTCTGCTGGCATTTATGGCAGAAAAGGCTTCCTCCCTCCCATACT 1660
Qy 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1661 CAGTTTGACTTGTGGAGCTCTGGATGGAAGAGGCTCTTAACCAAGCAAGGAGC 1720
Qy 61 LeuGluGlyThrProArgGlnSerArgGly 70
Db 1721 CTGGAGGCTCTCAGCGCCAAATCAGCGGCG 1750

RESULT 11
AF059485
LOCUS AF059485 9722 bp mRNA linear ROD 15-AUG-1998
DEFINITION Mus musculus Doc4 (Doc4) mRNA, complete cds.
ACCESSION AF059485
VERSION AF059485.1 GI:3170614

KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 9722)

Wang,X.-Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P.,
Zinszner,H. and Ron,D.
IDENTIFICATION OF NOVEL STRESS-INDUCED GENES DOWNSTREAM OF CHOP
EMBO J. 17 (13), 3619-3630 (1998)

MEDLINE 98315054
PUBMED 9649432
2 (bases 1 to 9722)

Wang,X.-Z. and Ron,D.
Direct Submission
Submitted (14-APR-1998) Skirball Institute, New York University
Medical Center, 550 First Ave., New York, NY 10016, USA

FEATURES
Location/Qualifiers
1..9722

source
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH/3T3"
/db_xref="taxon:10090"
/cell_line="NIH-3T3"
/note="tunicamycin-treated"
1..9722
/gene="Doc4"

/note="downstream of CHOP 4"

583..9060

CDS

/gene="Doc4"

/note="similar to Drosophila melanogaster team/odz and
human gamma-hergulin; type II transmembrane protein"

/codon_start=1

/product="DOC4"

/protein_id="AAC31807.1"

/db_xref="GI:3170615"

/translation="MDVKEKPYSLRTRDAERYTSSADSEBEGKGPQKSYSSSET
LKAYDQARLAYSRVXDMVPEAEFEETNFTLRELGLGEMTPPGTILYRTDIGL
PHCYSGMSADSLADLEADTVLSPEHVLWGRSTRSGSSCLSSRANSLTIDTDEHE
NETCAPLHCSSASSTIEQSPSPPPANESQRRLLNGVAOPTPDSDEEFPVPN
SFLVKSASISGVAANDHPSLONHRLRTPPLPHATPHQHAASINSINLRGNET
PRSNPSAPTSHLSGSEPPAGSAQEPHQAQDNVLSKIPIVETNKLQPLGWTQDN
LIEMDIFASRDGAYSQGHFFKPGGTSPLCTTSPGYLTSSTVYSPFPRPLRST
FSRPAFNLKPSKYCNWKAALSAILISATLIVLAYFAVHLFGLNHLWHPMEQMQ
MYEITEDTASSVPYTDVSLYSGGTGLTPDRKGKGAEPSPSLFEDPFIQSGEI
DVGRASQKIPGTFWRISOVFIDHPVHLKPNVSLKAALVIGYGRKGLPSPHTQIDFV
ELLDRLLTOEARLSEGPQRQSPVPSPSHETGFIQYLDGSIWHLAFYWDGKESEV
VSFLTALTESVNCPSNCGNGDCISGTCCHCFLGLGPDGCRASCPVLCSNGQYMKG
RCLCHGWKGAECVPTNQCIDVACSSHGTCTMGTCICNPGYKGESCEVDCMDPTCS
SRGVCVRGCHSCVGGTNCETPRATCDQCSGHGTFPLDTGLCNDPSTWGHDCSI
EICAADGGHGVCGVGTCTCRDGMGAACQDQACRCHPCAEHGTCTDCKECPGNGE
HCTTEGCPGLCNGRCCTDLNGWHVCQGLGRTGCTDCTMETGCKDNDGDLVD
CMDPDCLPCHVNPCLGSPDPLDIQTOAPVSOQLNSFYDRIKFLVGRDSTHS
IPGNPDGHCACVIRGOVMTSDGTPLGVNI SFINPLFGVTSRQDGLPVTNGG
ISILIRERAPPTQEHNLWLPDRFFVETIWMHENEIPSRDLSNFAFPNPVSP
SPLTSASSCAEKGPVPEIQALQEEIVAGCKMRSLYSLSKRTPGYKSVRSVLTHPT
IPFNLMVHLMVVEGRLEFRKFAAADLSYFIDKTDVYNQVFGSEAFVSILGYE
YESCPDGLTLEKRTAVLQGYRIDASLGGWSLDKHALNIOGSLHKGNGENFVSQ
PPVLSITMNGRRRSISCPNGLADGNKLAPVALTCGSDGSLVGVDFNIRFPFS
GNVTILMESHSPAKYLYLATDPMGAVLSDTNSRRVFKVKTSTTVKDLAKNSRVA
GTGQCLPDDTRCGDGKATEATLTPRGITVDKGLIYFVDTGTMIRVQNGLIST
LGSDNARPLSCDSVMEISQVRLMPTDLANPMONSLYLDNNVQLQISENHQV
RIVAGRMHCQVPGIDQIDLLSKVAIHATLESATALAVSHNGVLVIAETDEKKNIRQ
VTSSEISLVAGAPGCDKNDANCDGSDGYAKADAKLNTPLSPSLAYCAGDELYAD
LGNIRIRFKNKPFLNTQMYELSSPIDQELVLDFTSGHLYTQSLPTGYLYNFTY
TGDDI THITDNNGMVNVRRDTCMLPLVVPDQVYVWTGNTNSALRSTVTOGHEL
AMTYHNGSLLATKSNENGWTFEYVDSFGLNVTETPGQVSRFSDTSSVHVQV
ETSKDDVTITHLSSGAFYLLQDQVRNYYIAGDGLSLLLLANGHEVALQTEPHL
LAGTVNPGKRVNTPIDNGLNLEVRQKEQVTFVFGRLRVNHNRLLSLDFD
RVTEKIDHRRKFTLRLYDQGRSPFSPSRLNGVNTVSPGGHAGTQIGMS
ERMYDQAGRITSFADQKMSYTYLEKSNVHLHSQRYIFBPKNDRLSVTTPN
VAROTLETIRSVGYRNIYQPEGNASVIOQTEDGHLHLLFTYLTGTRVLYKYGLS
KLATLYDTTKVSTYDETAGMLATVNLQNEGFTCTIRYROI GLIDRQIREFTEGM
VNARPDYDNSFRVTSQAVINETPLPIDLYRYDDVSGKTEQKGFVYINDQII
TVAVTHSKHFDAYMRKEVOYETFRSLMYMTVQYDNGRVRVKKELKGVYANTRY
SYEDADGQLOQTVSINDKPLWRSYDNLGNLHLLSPGNSARLTLRLYDLDRITGLD
VQYKMDGSLRQRGDVFYNSAGLLIKAYNRASGWSRYRYDGLGRVRSKSHSH
HLQFPYADLNPNTKVTLYNHSSEITSLYDLOCHLFAVELSGDFEYIACDNIGTP
LAVPSGTGLMIKQILYAYGIBYMTNPNFQIILGYHGGYDPLTKLVHMGRRYDVL
AGRTSPDHELMKRLSSNIVPFLYKNNPISNSQDIKCFMTDVSMLLTFGOLF
HNVLPGKPKPTDAMEPSYELVHTQMTQWDSKSLGVQCEVQKQKAFVTLERED
QLYSTTSCQQAPETKKFASSGSIKGVKFKALDKGRVTVDIIISVANEGRRIAIL
NNARYLENLHFTIDGVTHYFVKPSEGLAILGLSGRRTLENGVNTVSQI
SGTRRTYDLOQFALCLNTRYGTVDDEKRVLELARAQVRAQWARQOORLREG
EGLRAWTDGKQOVLNTRGVQYDGFVTSVEQYPELSDSANNIHFMRQSEMR"

misc_feature

1813..1881

/note="putative; transmembrane-region site"

ORIGIN

Alignment Scores:

Pred. No.: 3,72e-31 Length: 9722
Score: 347.00 Matches: 67
Percent Similarity: 95.71% Conservative: 0
Best Local Similarity: 95.71% Mismatches: 3
Query Match: 93.28% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AF059485 (1-9722)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
REFERENCE
AUTHORS
Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,
Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
Millett, I. and Macdougall, J.R.
TITLE
JOURNAL
Proteins and nucleic acids encoding same
Patent: WO 02062999-A 39 15-AUG-2002;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
source
1..8473
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4,47e-20 Length: 8473
Score: 254.00 Matches: 48
Percent Similarity: 80.88% Conservative: 7
Best Local Similarity: 70.59% Mismatches: 13
Query Match: 68.28% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AX662357 (1-8473)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 1185 TTCTGGAGATCACAGCTCTTCATTGATCAGCACAGTTTCTTAAATTCATATCTCTT 1244

QY 22 GlyLysAlaAlaLeuValGlyLeuTyArgLysGlyLeuProSerHisThrGln 41
Db 1245 CAGAAGGATGATTGGAGTATATGCGCGAAAGGCTTACCGCTTCCCATCTCAG 1304

QY 42 PheAspPheValGluLeuLeuAspGlyArgLeuThrGlnGluAlaArgSerLeu 61
Db 1305 TATGACTTCGTGGAGCTCTGGATGGCAGAGCTGATGCCAGAGACGCGGAGCTG 1364

QY 62 GluGlyThrProArgGlnSerArg 69
Db 1365 CTTGAGACGGAGAGCGCGGCGG 1388

RESULT 14
AX662359
LOCUS AX662359 8487 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 41 from Patent WO02062999.
ACCESSION AX662359
VERSION AX662359.1 GI:29163219
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
REFERENCE
AUTHORS
Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,
Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
Millett, I. and Macdougall, J.R.
TITLE
JOURNAL
Proteins and nucleic acids encoding same
Patent: WO 02062999-A 41 15-AUG-2002;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
source
1..8487
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
REFERENCE
AUTHORS
Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,
Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
Millett, I. and Macdougall, J.R.
TITLE
JOURNAL
Proteins and nucleic acids encoding same
Patent: WO 02062999-A 39 15-AUG-2002;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
source
1..8473
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4,47e-20 Length: 8473
Score: 254.00 Matches: 48
Percent Similarity: 80.88% Conservative: 7
Best Local Similarity: 70.59% Mismatches: 13
Query Match: 68.28% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AX662359 (1-8487)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 1226 TTCTGGAGATCACAGCTCTTCATTGATCAGCACAGTTTCTTAAATTCATATCTCTT 1285

QY 22 GlyLysAlaAlaLeuValGlyLeuTyArgLysGlyLeuProSerHisThrGln 41
Db 1286 CAGAAGGATGATTGGAGTATATGCGCGAAAGGCTTACCGCTTCCCATCTCAG 1345

QY 42 PheAspPheValGluLeuLeuAspGlyArgLeuThrGlnGluAlaArgSerLeu 61
Db 1346 TATGACTTCGTGGAGCTCTGGATGGCAGAGCTGATGCCAGAGACGCGGAGCTG 1405

QY 62 GluGlyThrProArgGlnSerArg 69
Db 1406 CTTGAGACGGAGAGCGCGGCGG 1429

RESULT 15
AX662355
LOCUS AX662355 8645 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 37 from Patent WO02062999.
ACCESSION AX662355
VERSION AX662355.1 GI:29163217
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
REFERENCE
AUTHORS
Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,
Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
Millett, I. and Macdougall, J.R.
TITLE
JOURNAL
Proteins and nucleic acids encoding same
Patent: WO 02062999-A 37 15-AUG-2002;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
source
1..8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4,56e-20 Length: 8645
Score: 254.00 Matches: 48
Percent Similarity: 80.88% Conservative: 7
Best Local Similarity: 70.59% Mismatches: 13
Query Match: 68.28% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AX662355 (1-8645)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 1357 TTCTGGAGATCACAGCTCTTCATTGATCAGCACAGTTTCTTAAATTCATATCTCTT 1416

QY 22 GlyLysAlaAlaLeuValGlyLeuTyArgLysGlyLeuProSerHisThrGln 41
Db 1417 CAGAAGGATGATTGGAGTATATGCGCGAAAGGCTTACCGCTTCCCATCTCAG 1476

QY 42 PheAspPheValGluLeuLeuAspGlyArgLeuThrGlnGluAlaArgSerLeu 61

Db 1477 TATGACTTCGTGGAGCTCCCTGGATGGCAGCAGGCTGATTCGCCAGAGAGCAGCCGGAGCCTG 1536
QY 62 GluGlyThrProArgGlnSerArg 69
Db 1537 CTTGAGACGGAGAGAGCCGGGCGG 1560

Search completed: August 14, 2004, 11:51:25
Job time : 1270.76 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 21.0053 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_450_520
Perfect score: 372
Sequence: 1 TWRSGVFDHPVHLKFNVS.....RLLTQEARSLGTPRQSRGT 71

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NRM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @cgn_1_1_258 @runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO NMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 372 | 100.0 | 1680 | 3 | US-08-891-845-3 |
| 2 | 372 | 100.0 | 1680 | 4 | US-09-514-573-3 |
| 3 | 372 | 100.0 | 2387 | 3 | US-08-891-845-11 |
| 4 | 372 | 100.0 | 2387 | 4 | US-09-514-573-11 |
| 5 | 372 | 100.0 | 3111 | 3 | US-08-891-845-1 |
| 6 | 372 | 100.0 | 3111 | 4 | US-09-514-573-1 |
| 7 | 69 | 18.5 | 1083 | 4 | US-09-252-991A-11841 |
| 8 | 69 | 18.5 | 1407 | 4 | US-09-252-991A-11543 |
| 9 | 69 | 18.5 | 2304 | 4 | US-09-252-991A-11691 |
| 10 | 67.5 | 18.1 | 523 | 4 | US-09-669-751-81 |
| 11 | 66.5 | 17.9 | 1272 | 4 | US-09-252-991A-14082 |
| 12 | 66.5 | 17.9 | 1329 | 4 | US-09-252-991A-14036 |

| | | | | | |
|----|------|------|---------|---|----------------------|
| 13 | 66.5 | 17.9 | 1620 | 4 | US-09-252-991A-14186 |
| 14 | 66 | 17.7 | 3835 | 4 | US-09-566-921-120 |
| 15 | 65 | 17.5 | 945 | 4 | US-09-584-568C-3 |
| 16 | 65 | 17.5 | 1332 | 4 | US-09-584-568C-1 |
| 17 | 65 | 17.5 | 1426 | 4 | US-09-620-312D-967 |
| 18 | 63.5 | 17.1 | 513 | 4 | US-09-252-991A-1256 |
| 19 | 62.5 | 16.8 | 1233 | 4 | US-09-252-991A-3735 |
| 20 | 62.5 | 16.8 | 1491 | 4 | US-09-252-991A-3636 |
| 21 | 62.5 | 16.8 | 1812 | 4 | US-09-732-615-27 |
| 22 | 62.5 | 16.8 | 2242 | 3 | US-09-400-742-1 |
| 23 | 62.5 | 16.8 | 2242 | 3 | US-08-618-651A-1 |
| 24 | 62.5 | 16.8 | 2242 | 4 | US-09-215-252-1 |
| 25 | 62.5 | 16.8 | 2242 | 4 | US-09-370-989A-1 |
| 26 | 62.5 | 16.8 | 3301 | 4 | US-09-148-545-66 |
| 27 | 62.5 | 16.8 | 4884 | 4 | US-09-252-991A-3824 |
| 28 | 62 | 16.7 | 432 | 4 | US-08-956-171E-4061 |
| 29 | 62 | 16.7 | 480 | 4 | US-09-252-991A-5103 |
| 30 | 62 | 16.7 | 801 | 4 | US-09-252-991A-5129 |
| 31 | 62 | 16.7 | 1557 | 4 | US-09-252-991A-5090 |
| 32 | 62 | 16.7 | 2855 | 2 | US-08-852-153-1 |
| 33 | 62 | 16.7 | 3255 | 2 | US-08-852-153-3 |
| 34 | 62 | 16.7 | 3327 | 2 | US-08-852-153-3 |
| 35 | 62 | 16.7 | 22243 | 4 | US-08-956-171E-164 |
| 36 | 61.5 | 16.5 | 1080 | 4 | US-09-252-991A-15937 |
| 37 | 61.5 | 16.5 | 2721 | 4 | US-09-252-991A-16144 |
| 38 | 61.5 | 16.5 | 11601 | 2 | US-08-222-617A-3 |
| 39 | 61.5 | 16.5 | 11601 | 2 | US-08-222-617A-24 |
| 40 | 61 | 16.4 | 4403765 | 3 | US-09-103-840A-2 |
| 41 | 61 | 16.4 | 4411529 | 3 | US-09-103-840A-1 |
| 42 | 60.5 | 16.3 | 309 | 4 | US-09-252-991A-1165 |
| 43 | 60.5 | 16.3 | 1284 | 4 | US-09-252-991A-13863 |
| 44 | 60.5 | 16.3 | 1293 | 4 | US-09-252-991A-13668 |
| 45 | 60.5 | 16.3 | 1680 | 4 | US-09-252-991A-7909 |

ALIGNMENTS

RESULT 1
US-08-891-845-3
; Sequence 3, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1680 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-891-845-3

Alignment Scores:
Pred. No.: 1.3e-43 Length: 1680
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-08-891-845-3 (1-1680)

Qy 1 ThrPheTTPArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1348 ACTTCTCGAGATCTCAAGTGTTCATACACCATCTGTGCATCTCAATTCATGTGCT 1407
Qy 21 LeuGlyLysAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1408 CTGGAAAGGAGCCCTGTTGGCATTTATGGCAGAAAAGGCTCCCTCTTCACATACA 1467
Qy 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyArgLeuThrGlnGluAlaArgSer 60
Db 1468 CAGTTTGACTTTGTGAGCTGTGTGATGGCAGAGGCTCTTAACCCAGAGGCGCGGAGC 1527
Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1528 CTAGAGGGAGCCCGCGCCAGTCTCGGGGAAC 1560

RESULT 2

US-09-514-573-3
Sequence 3, Application US/09514573
Patent No. 6500941
GENERAL INFORMATION:
APPLICANT: Schaefer, Gabriele M.
APPLICANT: Sliwowski, Mark
TITLE OF INVENTION: Gamma-Herregulin
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/514,573
FILING DATE: 28 FEB 2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/891845
FILING DATE: 10 JULY 1997
APPLICATION NUMBER: 60/021640
FILING DATE: 07/12/96
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P1043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 1680 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-09-514-573-3

Alignment Scores:
Pred. No.: 1.3e-43 Length: 1680
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-09-514-573-3 (1-1680)

Qy 1 ThrPheTTPArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1348 ACTTCTCGAGATCTCAAGTGTTCATACACCATCTGTGCATCTCAATTCATGTGCT 1407
Qy 21 LeuGlyLysAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1408 CTGGAAAGGAGCCCTGTTGGCATTTATGGCAGAAAAGGCTCCCTCTTCACATACA 1467
Qy 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyArgLeuThrGlnGluAlaArgSer 60
Db 1468 CAGTTTGACTTTGTGAGCTGTGTGATGGCAGAGGCTCTTAACCCAGAGGCGCGGAGC 1527
Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1528 CTAGAGGGAGCCCGCGCCAGTCTCGGGGAAC 1560

RESULT 3

US-08-891-845-11
Sequence 11, Application US/08891845
Patent No. 6096873
GENERAL INFORMATION:
APPLICANT: Schaefer, Gabriele M.
APPLICANT: Sliwowski, Mark
TITLE OF INVENTION: Gamma-Herregulin
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/891,845
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/021640
FILING DATE: 07/12/96
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P1043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2387 base pairs
TYPE: Nucleic Acid

```
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-891-845-11

Alignment Scores:
Pred. No.: 2,1e-43 Length: 2387
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-08-891-845-11 (1-2387)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 853 ACTTTCGGAGATCTCAAGTGTTTCATAGACCATCTCTGTCATCTGAAATTCATGTGTCT 912
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 913 CTGGGAAAGCAGCCCTGGTGGCATTTATGGCAGAAAGGCCCTCCCTTCACATACA 972
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyArgLeuThrGlnGluAlaArgSer 60
Db 973 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCTTACCCAGGAGGCGCGGAGC 1032
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1033 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1065

RESULT 4
US-09-514-573-11
; Sequence 11, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
```

```
; TOPOLOGY: Linear
; US-09-514-573-11

Alignment Scores:
Pred. No.: 2,1e-43 Length: 2387
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-09-514-573-11 (1-2387)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 853 ACTTTCGGAGATCTCAAGTGTTTCATAGACCATCTCTGTCATCTGAAATTCATGTGTCT 912
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 913 CTGGGAAAGCAGCCCTGGTGGCATTTATGGCAGAAAGGCCCTCCCTTCACATACA 972
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyArgLeuThrGlnGluAlaArgSer 60
Db 973 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCTTACCCAGGAGGCGCGGAGC 1032
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1033 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1065

RESULT 5
US-08-891-845-1
; Sequence 1, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-891-845-1
```

Alignment Scores:

Pred. No.: 3 01e-43 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-08-891-845-1 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTCGAGATCTCAAGTGTTCATAGACCATCTCTGTGCATCTGAAATTCATGTGCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrglyArgLysGlyLeuProSerHisThr 40
Db 1741 CTGGAAAGGCGACCTGCTGGCATTTATGCGAGAAAGGCTCCCTTCATACATA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGAGCTGCTGATGGCAGAGGCTCCTAACCCAGAGGCGCGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCGCCAGTCTCGGGAACT 1893

RESULT 6

US-09-514-573-1
; Sequence 1, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinFatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-09-514-573-1

Alignment Scores:

Pred. No.: 3 01e-43 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-09-514-573-1 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTCGAGATCTCAAGTGTTCATAGACCATCTCTGTGCATCTGAAATTCATGTGCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrglyArgLysGlyLeuProSerHisThr 40
Db 1741 CTGGAAAGGCGACCTGCTGGCATTTATGCGAGAAAGGCTCCCTTCATACATA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGAGCTGCTGATGGCAGAGGCTCCTAACCCAGAGGCGCGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCGCCAGTCTCGGGAACT 1893

RESULT 7

US-09-252-991A-11841/c
; Sequence 11841, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11841
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11841

Alignment Scores:

Pred. No.: 1.87 Length: 1083
Score: 69.00 Matches: 28
Percent Similarity: 44.00% Conservative: 5
Best Local Similarity: 37.33% Mismatches: 30
Query Match: 18.55% Indels: 12
DB: 4 Gaps: 4

US-10-029-020-14_COPY_450_520 (1-71) x US-09-252-991A-11841 (1-1083)

QY 9 IleAspHisProValHisLeuLysPheAsnValSerLeuGlyLysAlaAlaLeuVal--- 27
Db 1081 ATTGATCACACTGGCCACCTCCAGCATCCAGCTCGGCTGGGCGTACTTGTGCGC 1022

QY 28 GlyIleTyrglyArgLysGlyLeuProPro-----SerHisThrGln 41
Db 1021 GGCTTGGGGTCTGCTGCCGACACCGCCGCCCTCGACAGCGCGAGTGCAGCGAG 962

QY 42 PheAsp-----PheValGluLeuLeuAspGly-----ArgArgLeuThrGln 56
Db 961 CCCGACACTGCGGTGCTGCTATTTCTGCTAGTGTCTTCGGAAGCGAGTGGTCTT 902

QY 57 GluAlaArgSerLeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 901 GCTGCTGCTGCGTAGAGGCTGAGCTGGGCAAGTGGCCCGGACG 857

Percent Similarity: 44.00% Conservative: 5
Best Local Similarity: 37.33% Mismatches: 30
Query Match: 18.55% Indels: 12
DB: 4 Gaps: 4

US-10-029-020-14_COPY_450_520 (1-71) x US-09-252-991A-11691 (1-2304)

QY 9 lIeAspHisProValHisLeuLysPheAenValSerLeuGlyLysAlaLeuVal--- 27
Db 1977 ATTGATCACACTGGCCACCTGGGCATCCAGCTCGGCTGGGCGTGGGCGTACTTGTGCGC 1918
QY 28 GlyIleTyGlyArgLysGlyLeuPro-----SerHisThrGln 41
Db 1917 GGCTCGCGGGTCGCTGCCGACACCCCGCCCTCGAACAGCGGCGAGGCTGAGACGCAG 1858
QY 42 PheAsp-----PheValGluLeuLeuAspGly-----ArgArgLeuLeuThrGln 56
Db 1857 GCGCACACTGTCGGTGTGATTTCTGCTCGTAGTGCTTTCCGAAGCCCGAGTGCCTT 1758
QY 57 GluAlaArgSerLeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1797 GCTGCTGCTGGCTAGAGGTGAGCCTGGGCAAGTGGCCGCGCAG 1753

RESULT 10
US-09-669-751-81
; Sequence 81, Application US/09669751
; Patent No. 6551575
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Ralph J.
; TITLE OF INVENTION: Methods for Identifying Compounds for
; TITLE OF INVENTION: Motion Sickness, Vertigo and Other Disorders Related to
; TITLE OF INVENTION: Balance and the Perception of Gravity
; FILE REFERENCE: P-NI 3864
; CURRENT APPLICATION NUMBER: US/09/669,751
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 60/168,579
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Drosophila
US-09-669-751-81

Alignment Scores:
Pred. No.: 1,14 Length: 523
Score: 67.50 Matches: 19
Percent Similarity: 56.52% Conservative: 7
Best Local Similarity: 41.30% Mismatches: 17
Query Match: 18.15% Indels: 3
DB: 4 Gaps: 2

US-10-029-020-14_COPY_450_520 (1-71) x US-09-669-751-81 (1-523)

QY 23 LysAlaAlaLeuValGlyIleTyGlyArgLysGlyLeuProSerHisThrGlnPhe 42
Db 340 AAGGGTGTGTGGTGTGGTGTATACCAAGGATGGCGATAAGCATCGAAACACCGCA 399
QY 43 AspPheValGluLeuLeuAsp-----GlyArgArgLeuLeuThrGln--GluAlaArg 59
Db 400 AATCAGTGCACCTTGGATGATGCTCTGGTGGCAAGCTGTTGACCCCTGATCCGTGAACGT 459
QY 60 SerLeuGluGlyThrPro 65
Db 460 GGAATGGACGGCACTCCC 477

RESULT 11
US-09-252-991A-14082/c
; Sequence 14082, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14082
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14082

Alignment Scores:
Pred. No.: 5.3 Length: 1272
Score: 66.50 Matches: 20
Percent Similarity: 50.00% Conservative: 2
Best Local Similarity: 45.45% Mismatches: 19
Query Match: 17.88% Indels: 3
DB: 4 Gaps: 2

US-10-029-020-14_COPY_450_520 (1-71) x US-09-252-991A-14082 (1-1272)

QY 10 AspHisProValHisLeuLysPheAsnValSerLeuGlyLysAlaLeuValGlyIle 29
Db 681 GATCACCGCGGCACCTCTGTTCCGGGTGCTGTGCTGCGGAACTCGCGGATA 622
QY 30 TyrGlyArgLysGlyLeuProSerHisThrGlnPheAspPheValGlu-----Leu 47
Db 621 CAGCGCGCGCAGGGTCACCCGCCGAT---ACCGCGCGCGCCAGTCGAGCAGGACCTT 565
QY 48 LeuAspGlyArg 51
Db 564 CTCGATGGCCGG 553

RESULT 12

US-09-252-991A-14036/c
; Sequence 14036, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14036
; LENGTH: 1329
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14036

Alignment Scores:
Pred. No.: 5.63 Length: 1329
Score: 66.50 Matches: 20
Percent Similarity: 50.00% Conservative: 2
Best Local Similarity: 45.45% Mismatches: 19
Query Match: 17.88% Indels: 3
DB: 4 Gaps: 2

US-10-029-020-14_COPY_450_520 (1-71) x US-09-252-991A-14036 (1-1329)

QY 10 AspHisProValHisLeuLysPheAsnValSerLeuGlyLysAlaLeuValGlyIle 29
Db 724 GATCACCGCGGCACCTCTGTTCCGGGTGCTGTGCTGCGGAACTCGCGGATA 665

QY 30 TyrGlyArgLysGlyLeuProSerHisThrGlnPheAspPheValGlu-----Leu 47
Db 664 CAGCGCGCGCAGGGTCACCCGCCGAT---ACCGCGCGCGCCAGTCGAGCAGGACCTT 608
QY 48 LeuAspGlyArg 51
Db 607 CTCGATGGCCGG 596

RESULT 13

US-09-252-991A-14186
; Sequence 14186, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14186
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14186

Alignment Scores:
Pred. No.: 7.37 Length: 1620
Score: 66.50 Matches: 20
Percent Similarity: 50.00% Conservative: 2
Best Local Similarity: 45.45% Mismatches: 19
Query Match: 17.88% Indels: 3
DB: 4 Gaps: 2

US-10-029-020-14_COPY_450_520 (1-71) x US-09-252-991A-14186 (1-1620)

QY 10 AspHisProValHisLeuLysPheAsnValSerLeuGlyLysAlaLeuValGlyIle 29
Db 1183 GATCACCGCGGCACCTCTGTTCCGGGTGCTGTGCTGCGGAACTCGCGGATA 1242
QY 30 TyrGlyArgLysGlyLeuProSerHisThrGlnPheAspPheValGlu-----Leu 47
Db 1243 CAGCGCGCGCAGGGTCACCCGCCGAT---ACCGCGCGCGCCAGTCGAGCAGGACCTT 1299
QY 48 LeuAspGlyArg 51
Db 1300 CTCGATGGCCGG 1311

RESULT 14

US-09-566-921-120/c
; Sequence 120, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 120
; LENGTH: 3835
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No. 6682888 196545.7
; NAME/KEY: unsure
; LOCATION: 3743-3768
; OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-120

Alignment Scores:
Pred. No.: 28 Length: 3835
Score: 66.00 Matches: 19
Percent Similarity: 40.00% Conservative: 11
Best Local Similarity: 25.33% Mismatches: 25
Query Match: 17.74% Indels: 20
DB: 4 Gaps: 2

US-10-029-020-14_COPY_450_520 (1-71) x US-09-566-921-120 (1-3835)

QY 11 HisProValHisLeu----- 15
Db 2940 CACCCCTACTCACTTACACCTGCTCTAGACAAAGTTGTTCAATGACTTAAAAAATACTT 2881
QY 16 ---LysPheAsnValSerLeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGly 34
Db 2880 TTTCAATTCCCTTTGGCATAACCACTGCTGGACTGCTGGGTCACATGGA----- 2830
QY 35 LeuProProSerHisThrGlnPheAspPheValGluLeuLeuAspGlyArgLeuLeu 54
Db 2829 ---CCTACGATTTCCTTCCACATGCCCTTTATTGATCTGGGGTCTGGTAAGTTGATGCTA 2773
QY 55 ThrGlnGluAlaArgSerLeuGluGlyThrProArgGlnSerArg 69
Db 2772 ATGCAGAGCAATAGTAGAGTCCAGGCACTAAATATGTCAGCAAG 2728

RESULT 15

US-09-584-568C-3
; Sequence 3, Application US/09584568C
; Patent No. 6500657
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria, Alexandra et al.
; TITLE OF INVENTION: 33167, A NOVEL HUMAN HYDROLASE AND USES THEREFOR
; FILE REFERENCE: MNI-140
; CURRENT APPLICATION NUMBER: US/09/584,568C
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 60/193,954
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(945)
US-09-584-568C-3

Alignment Scores:
Pred. No.: 5.81 Length: 945
Score: 65.00 Matches: 19
Percent Similarity: 53.45% Conservative: 12
Best Local Similarity: 32.76% Mismatches: 23
Query Match: 17.47% Indels: 4
DB: 4 Gaps: 1

US-10-029-020-14_COPY_450_520 (1-71) x US-09-584-568C-3 (1-945)

QY 12 ProValHisLeuLysPheAsnValSerLeuGlyLysAlaAlaLeuValGlyIleTyrGly 31
Db 154 CCGCTTCGCTTCTCCTACAGGCTTCTGGACGGGGAGGAGCCCTCCCGCGCGTCTTT 213
QY 32 ArgLysGlyLeuProProSerHisThrGlnPheAspPheValGluLeuLeu----- 48
Db 214 TTGCACGGGCTCTTCGGCAGCAAAACTTAACCTCACTCCATCGCCCAAGATCTTGGCCAG 273

QY 49 ---AspGlyArgArgLeuLeuThrGlnGluAlaAspSerLeuGluGlyThrPro 65
Db 274 CAGACAGGCCGTAGGTGCTGACGGTGGATGCTCGTAACCAACCGTGCACAGCCCC 327

Search completed: August 14, 2004, 19:39:45
Job time : 37.0053 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: August 14, 2004, 02:35:12 ; Search time 534 Seconds
(without alignments) 3182.171 Million cell up

Title: SEQ14-X-AT-28-64-76

Perfect score:

Sequence: 1 MDVKERKPYRSLTRRRDAER.....EITEDTASSWPVPTDVSLYP 400

Scoring table: BLOSUM62

| |
|---|
| <p> Xgapop 10.0 , Xgapext 0.5 Ygapop 10.0 , Ygapext 0.5 Fgapop 6.0 , Fgapext 7.0 Delop 6.0 , Delext 7.0 </p> |
|---|

Searched: 3373863 seqs. 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlh
```

-Q=/cgn2_1/USPTO_spool/MITRA020/runat 06082004 114100 196/app query.fasta 1 583

```
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
```

```
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
```

```
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
```

```
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
```

```
-USER=MITRA020 @CGN_1_1_470 @runat 06082004 114100 196 -NCPU=6 -ICPU=3
```

```
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
```

```
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
```

```
-FGAPEXT=/ -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : N Geneset 29-Jan-04.*

```
Database : N_Geneseq_29Jan04:**
1:  geneseq1980s:**
```

1: geneseq11980s.*
2: geneseq11990s.*

2: geneseqnr330s:*

4: geneseqn2001as:*

1: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| Result No. | Score | Query % | | Length | DB | ID | Description |
|------------|-------|---------|--------|--------|----------|-----------|-------------|
| | | Match | Length | | | | |
| 1 | 2127 | 99.8 | 8354 | 6 | ABS52100 | Human | TEN |
| 2 | 2122 | 99.5 | 3111 | 2 | Aav19251 | Human | gam |
| 3 | 2122 | 99.5 | 3111 | 2 | AAX87705 | Gamma-her | |
| 4 | 2122 | 99.5 | 3111 | 3 | AAD00791 | Human | Her |
| 5 | 2122 | 99.5 | 3111 | 4 | AAS18526 | DNA | encod |
| 6 | 2082 | 97.7 | 8438 | 6 | ABN85378 | NOV | |
| 7 | 2066 | 96.9 | 8645 | 6 | Abs78652 | Human | cdN |
| 8 | 1285 | 60.3 | 2387 | 2 | Aav19252 | Human | gam |

```

FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /*standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /*tag= c
FT /*standard_name= "Single nucleotide polymorphism (SNP)"
FX
FX WO200257453-A2.
FX
FX 25-JUL-2002.
FX
FX 19-DEC-2001; 2001WO-US050331.
FX
FX 19-DEC-2000; 2000US-0256704P.
FX PR 20-DEC-2000; 2000US-0257314P.
FX PR 02-MAY-2001; 2001US-0288153P.
FX PR 29-MAY-2001; 2001US-0294075P.
FX PR 24-JUL-2001; 2001US-0307508P.
FX PR 10-AUG-2001; 2001US-0311590P.
FX PR 10-AUG-2001; 2001US-0311613P.
FX PR 29-AUG-2001; 2001US-0315617P.
FX PR 14-SEP-2001; 2001US-0322359P.
FX
FX (CURA-) CURAGEN CORP.
FX
FX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
FX Stone DJ, Anderson D, Shinkets RA, Burgess CB, Zexhusen BD, Liu X;
FX Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
FX
FX WPI; 2002-590744/63.
FX DR P-PSDB; ABG70388.
FX
FX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
FX atherosclerosis, metabolic disorders, diabetes, obesity, infectious
FX disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
FX cancer.
FX
FX Claim 8; Page 50-52; 318pp; English.
FX
FX The present invention relates to new NOVX polypeptides. The invention is
FX useful for treating or preventing a NOVX-associated disorder such as
FX cardiomyopathy or atherosclerosis, where the disorder is related to cell
FX signal processing and metabolic pathway modulation in a subject.
FX
FX disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
FX disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
FX disorders, haematopoietic disorders and various cancers. The molecules of
FX the invention are also useful for treating or preventing cirrhosis,
FX pancreatitis, learning and memory defects, infertility, congenital heart
FX defects, acne, hair growth, pigmentation disorders, endocrine disorders,
FX respiratory disease, gastro-intestinal diseases, reproductive, health,
FX allergic and inflammation, bone marrow transplantation, urinary system
FX disorders, neuropsychiatric disorders and age-related disorders. The
FX present nucleic acid sequence represents a NOVX gene. This sequence
FX encodes a NOVX protein of the invention
FX
FX Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;
FX
FX
FX Alignment Scores:
FX Pred. No.: 6 54e-127 Length: 8354
FX Score: 2127.00 Matches: 397
FX Percent Similarity: 99.25% Conservative: 0
FX Best Local Similarity: 99.25% Mismatches: 3
FX Query Match: 99.77% Indels: 0
FX DB: 6 Gaps: 0
FX
FX SEQ14-X-AT-28-64-76 (1-400) x ABS52100 (1-8354)
FX
FX 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
FX
FX 35 ATGGACGTGAAGGAGAGAGCCCTTACCGCTCGTGACCCCGGCCCGGAGCGGC 94

```

```

QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB CGCTACACCACTCGTCCGCGGACAGGAGGAGGAAAGCCCGCAGAAATCGTACAGC 154
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB TCCAGGAGAGCCCTGAAGGCTTACAGCAGGAGCCCGCTAGCCTATGGCAGCGCGCTC 214
QY 61 LysAspPhe***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB AAGGACATTGTCCCGCAGGAGGCGGAGAAATTCGCCGACAGGTGCCAACTTCACCTG 274
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspPhe 100
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB CGGAGCTGGGGCTGGAAGAGTAAACGCCCTCAGCGGAGCCCTGTACCGGACAGACATT 334
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB GGCCTGCCCAATGCGGCTACTCCATGGGGGCTGCTCTGTATGCCGACATGGAGGCTGAC 394
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB ACGGTGCTGTCCCTTGAGCACCCCGTGGCTCTGTGGGGCCGAGGACACAGGTCCAGGCGC 454
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB AGCTCTCGCTCTCCAGCGCGGCCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 514
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB AACACTGAGACTGATCATCCGGCGGCGCTGCAAGAACCCAGCGCGGCTCCGGAGCGCGCG 574
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaSerIleAsnSerLeu 200
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB CGCGCGCTCTCCAGCGGCCACACCCCAACAGCAGCACCGCGGCTCCATTAATCCCTG 634
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB AACCGGGCACTTTCAGCGCGGAGGAGCAACCCAGCGCGGCCCGCCCGACCGACCATCGCTC 694
QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB TCCGAGAGCGCCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 754
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB CTCACAGCAACATCCCTCTGGAGACCAAGAACCTAGGCAAGCAGGACCAATCTCTAGGACA 814
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB TTGACGAGCAACCTCATTGAGATGGACATTCGCGCGCTCCCGCCATGATGGGGCTTAC 874
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB AGTGACGCGCACTTCCTCTTCAAGCTGGAGGAGCCCTCCCGCTCTCTGACCAACATCA 934
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB CAGAGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCCGCCCGACCGCTGCCCGCG 994
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysPheProSerLysTyrCysAsnTrpLys 340
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB AGCACCTTCGCCCGCGCGCGCTTTAAACCTCAAGAGCCCTCCAGTACTGTAACTGGAAG 1054
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB TGGCGAGCCCTGAGCGCATCTGTCATCTAGCCACTCTGTGTCATCTCTGCTGGCATCTTT 1114
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB GTGGCCATGCACCTGTTGGCTTAACTGGCACCTGAGCGCGATGGAGGGGCGAGATGTAT 1174

```

QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1175 GAGATCAGGAGGACACAGCCAGTGGCTGTGCTGCCAACCGACGCTCTCCCTATACCCC 1234

RESULT 2

AAV19251
ID AAV19251 standard; cDNA; 3111 BP.

XX AAV19251;

DT 17-AUG-1998 (first entry)

XX Human gamma-hergulin cDNA.

XX Gamma-hergulin; gamma-HRG; human; autocrine growth factor;
KW breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor;
KW cell proliferation; cell differentiation; cell survival;
KW neurological disorder; muscular disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 534..2645
ET CDS /*tag= a

XX WO9802541-A1.

XX 22-JAN-1998.

XX 08-JUL-1997; 97WO-US011841.

XX 12-JUL-1996; 96US-0021640P.

XX (GETH) GENENTECH INC.

XX Schaefer GM, Sliwkowski M;

XX WPI; 1998-110589/10.

DR P-PSDB; AAW44817.

XX DNA encoding gamma-hergulin - used to activate ErbB receptor and to
enhance proliferation, differentiation or survival of a cell.

XX Claim 21; Fig 1A-C; 81pp; English.

CC This nucleic acid molecule codes for human gamma-hergulin (gamma-HRG)
CC (see AAW44817), a novel member of the heregulin superfamily, that has a
CC unique N-terminal domain not present in previously identified heregulins.
CC Gamma-HRG is secreted by human breast cancer MDA-MB-175 cells. It leads
CC to the formation of a constitutive active receptor complex and stimulates
CC the growth of these cells in an autocrine manner. The nucleic acid was
CC isolated by screening a MDA-MB-175 with a probe corresponding to the EGF-
CC like domain and part of the N-terminal sequences of HRG-beta-3. It can be
CC used for the recombinant production of gamma-HRG, or for in vivo ex
CC vivo gene therapy. A claimed nucleic acid, which is complementary to the
CC nucleic acid sequence encoding the N-terminal domain of gamma-HRG, is
CC able to reduce production of gamma-HRG by MDA-MB-175 cells. The invention
CC provides claimed methods for activating an ErbB receptor and for
CC enhancing proliferation, differentiation or survival of a cell by
CC contacting the cell (preferably a glial or muscle cell) with gamma-HRG

XX Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,42e-127 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 2 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x AAV19251 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGAGCTGAAGGAGGAGGAGGCTTACCGCTCGCTGACCGCGCGCGCGAGCGC 393
QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCACTCGTCCCGGAGACAGCGAGGAGGGCAAGAGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCTGAAGGCTTACGACAGGAGCGCCGCTAGCCTATGGCAGCCGCGTC 513
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 514 AAGGACATTTGTCCGCGAGGAGCCGAGGAATTTCTGCCGACAGGTGCCCACTTCAACCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CCGGAGCTGGGGCTGGAGAGTAACGCCCTCTACGGGACCTGTACCGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCTCCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 694 ACGTGTCTGTCCCTCAGCACCCCGTGTGTGGGCGGAGCACACGGTCAGGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTGTCTGTCCAGCGGGCCCAATTCATATCTACACTCACCAGACCCAGGATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGCGGCTGCAGAACCCAGCGGGCTCCGGAGCGCGCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGCTCTCGACGCCCCACACCCCCACACAGCACACCGCGGCTCTCACTTAACCTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCAGCGCGAGAGCAACCCAGCGCGCGCCCGACGAGCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGAGAGAGCCCTTCCCGCGCGCGCGCGAGAGCTGCCACCGCCAGGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCTTGGAGACCCAGAAACCTTAGCAAGCAGCATTCTTAGGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGAGGACCACTCATTTGAGATGACATTCGCGGCTCCCGCATGATGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCTCTTCAAGCTTGAGGACCTTCCCGCTCTTCTGCACCATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGACTCTCTCCGCGCGCGCGCGCTGCCCCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCAGCTTCGCGCGCGCGCGCTTTTAACTCAAGAGCCCTCCAGTACTGTAACTGGAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1354 TGGCAGCGCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCTATCTCTGCTGTCATAC 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

Db 1414 GTGGCCATGCACCTGTTGGCTAACTGGACCTGCAGCGGATGGAGGGGCACATGTAT 1473

Qy 381 GluIleThrGluAspThrAlaSerSerTTPProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCAGTGGCTGTGCCAACCGACGCTCCCTATACCCC 1533

RESULT 3

AAX87705

ID AAX87705 standard; cDNA; 3111 BP.

XX AC AAX87705;

DT 26-OCT-1999 (first entry)

DE Gamma-herregulin cDNA.

XX Gamma-herregulin; HRG-alpha; human; ligand; HER2; HER3; HER4; receptor;
XX lung surfactant; respiratory distress syndrome; emphysema;
XX epithelial growth factor; therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 334..2640

XX FT /*tag= a

XX WO9939729-A2.

XX 12-AUG-1999.

XX 03-FEB-1999; 99WO-US002390.

XX 04-FEB-1998; 98US-00020598.

XX (GETH) GENENTECH INC.

XX (IOWA) UNIV IOWA RES FOUND.

XX Sliwkowski M, Kern JA;

XX WPI; 1999-494213/41.

XX DR P-ESDB; AAY06639.

XX Heregulin ligands can be used to induce epithelial cell growth, and to
XX promote repair and healing of tissue damage or injury.

XX Disclosure; Page 111-116; 120pp; English.

XX This is the nucleotide sequence of cDNA coding for gamma-herregulin (gamma
XX -HRG, see AAY06639). The invention provides HRG ligands, including gamma-
XX HRG, that have affinity for and stimulate HER2, HER3 and/or HER4
XX receptors in autophosphorylation. A new method of treating respiratory
XX distress syndrome in humans uses HER2, HER3 and/or HER4 receptor ligands
XX as epithelial growth factors. A novel method of inducing epithelial cell
XX growth and/or proliferation comprises contacting a normal epithelial cell
XX which expresses HER2, HER3 and/or HER4 receptors with an isolated ligand
XX which activates HER2, HER3, HER4 receptors or their combination. Also
XX claimed are methods of increasing lung surfactant protein A, or of
XX treating chronic obstructive pulmonary disease, respiratory distress or
XX emphysema, by administering an effective amount of an isolated HER ligand
XX to a patient

SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 4.42e-127 | Length: | 3111 |
| Score: | 2122.00 | Matches: | 396 |
| Percent Similarity: | 99.00% | Conservative: | 0 |
| Best Local Similarity: | 99.00% | Mismatches: | 4 |
| Query Match: | 99.53% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

SEQ14-X-AT-28-64-76 (1-400) x AAX87705 (1-3111)

Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGACGTGAAGAGAGAGAGACCTTACCGCTCGCTGACCGCGCCGCGAGCGCGC 393
Qy 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCACTCGTCCGCGGACAGCGAGAGGGCAAGCCCGCAGAAATCGTACAGC 453
Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaGluLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCCCTGAAGGGCTTACGACCGAGACCCCGCTAGCTTATGGCAGCGCGT 513
Qy 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 514 AAGGACATTTGCGCGCAGGAGGCGCAGAAATCTGCGCGCACAGGTGCCAATTCACCTG 573
Qy 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCTCACGGGACCCCTGTACCGGACAGACATT 633
Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCCTCCCCCACTGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
Qy 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCTCTGAGCACCCCGCTGCTGTGGGGCGGAGACACACGGTCAGGGCGC 753
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTCGCTGTCCAGCGGGCCAATTCCAAATCTCACACTCACCGACACCGAGCATGAA 813
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGCGGCTGTGAGAACCCACCGCGGCTCCGACGCGCGCG 873
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGGCTCTCGCACGCCCCACACCCCAACAGCACCGCGGCTCCATTAATCTCCCTG 933
Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGGCACTTTCAGCGCGAGAGCAACCCAGCGCGCGCCCGCCAGCGACCTCGCTC 993
Qy 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGAGAGCCCCCTGCGGGCGCGCCAGGAGCTGCCACGCCCGCCAGGAGAACTGGCTG 1053
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCTCTGGAGACCCAGAAACCTAGGCAAGAGCCATTCTTAGGGACA 1113
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGAGGACAACTTCATTGATGATGAGACATTTCTCGCGCGCTCCCGCCATGATGGGGCTTAC 1173
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCTCTTCAAGCCTGGAGGACCTCCCGCGCTCTTCTGGACCAATCA 1233
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCGCGCCCGACCCCTGCCCGC 1293
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCCTTGGCGCGCGCGCTTTAACTCAAGAGAGCCCTCCAAAGTACTGTAACCTGGAG 1353
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCAGCCCTGAGGCGCATGTCATCTCAGCCACTCTGGTCATCTCTCTGGCATACCTT 1413

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTTGGCCCTAAACTGGCAGCTGCAGCCGATGGAGGGCGAGATGTAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGAGTGGCCCTGTGCCAACCGAGCTCTCCCTATACCCC 1533

RESULT 4

AAAD00791
ID AAAD00791 standard; cDNA; 3111 BP.
XX
AC AAAD00791;
XX 21-SEP-2000 (first entry)
DT
DE Human Heregulin variant, gamma-HRG encoding cDNA.
XX
KW Heregulin; variant; gamma-HRG; human; inner-ear-supporting cell;
KW activator; HER2 receptor; HER3 receptor; HER-2 oncogene; chromosome 8p;
KW hair cell related hearing disorder; ototoxic injury; tissue damage;
KW acoustic assault; degenerative hearing loss; balance impairment;
KW treatment; hair cell; surgical injury; physical injury;
KW inner ear disorder; ss.
XX

OS Homo sapiens.

XX
FH Key Location/Qualifiers
FT CDS 334..2640
FT /tag= a
FT /product= "Human Heregulin variant, gamma-HRG protein"

XX WO200027426-A1.

XX 18-MAY-2000.

XX 28-OCT-1999; 99WO-US025744.

XX 07-NOV-1998; 98US-0107522P.

XX (GETH) GENENTECH INC.

XX Gao W;

XX WPI; 2000-376313/32.

XX P-PSDB; AAY71176.

XX

PT Method for inducing hair cell generation and inner-ear-supporting cell
PT growth regeneration and proliferation, useful for treating hearing
PT disorders.

XX Disclosure; Fig 7; 141pp; English.

XX The patent discloses a method for inducing hair cell generation, or inner
CC -ear-supporting cell growth, regeneration, and/or proliferation, by
CC heregulins (HRG), the ligand for HR2/HER3 dimeric receptors. Heregulin
CC proteins function as activators of HER-2 oncogene and result from
CC alternate splicing of a single gene mapped to chromosome 8p. The two
CC major types, alpha and beta HRG's are based on two variant EGF-like
CC (epidermal growth factor) domains, that differ in their C-terminal ends.
CC This method can be used to increase the number of inner-ear-supporting
CC cells and for treatment of hair cell related hearing disorders and
CC disease states associated with tissue damage, e.g. ototoxic injury,
CC acoustic assault, degenerative hearing loss, balance impairments, damage
CC associated with surgery or physical injury and inner ear disorders
CC related to hair cell dysfunction. The present sequence is the human
CC heregulin variant, gamma-HRG protein encoding cDNA, isolated from a
CC native HRG source, such as human MDA-MB-175 cells. It has atleast one
CC biological property as the native sequence

XX Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

SQ Alignment Scores:

Pred. No.: 4,42e-127 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 3 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x AAD00791 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGAGCTGAAGGAGAGAGAGCTTACCGCTGCTGACCCGGCGCCGAGCCGAGCGC 393
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCAGCTCGTCCGCGACACGAGAGGGGCAAGAGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCAGCGAGACCCCTGAAGGCCTACGACCAAGAGCGCCCGCTAGCCTATGCGAGCGCGTC 513
QY 61 LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db 514 AAGGACATTGTCCCGCAGGAGGCGGAGAAATTCTGCGCACAGTGCACACTTCACCCCTG 573
QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGGAGCTGGGGCTGGAAGAGTAACGCCCGCTCAGCGACCCCTGTACCGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCCACTGGCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATGGAGGGTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 694 ACGTGTCTGTCCCTGAGCACCCCGTGTCTGTGGGGCGGAGCACAGGTTCAGGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTGCTGTCCAGCGGGGCCAATTCCAAATCTCACACTCACCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGCGGCGCTGCAAGAACACGCGCGGCTCCGGACCGCGCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaSerIleAsnSerLeu 200
Db 874 CCGCCGCTCTCGACGCGCCACACCCCAACACGAGCACCGCGGCTCCATTAACTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGGCAACTTCACGCGGAGAGCAACCCGAGCCCGCGCCCGCCACCGACCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCGGAGAGCCCCCTGCGCGCGGCGCCAGAGACCTGCCCCAGCCCGAGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCCCTCGAGACCAAGAACTAGGCAAGACGCCATTCTCAGGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGAGGAGCAACCTATTGAGATGACATTCTCGGCGGCTCCCGCATGATGGGGTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGAGGGGCACTTCTCTTCAAGGCTTGAGGACACTCCCGGCTCTCTCGCCACATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGATACTCTCTCCGCGCCGACCCCTGCCCGC 1293


```
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCACCGGAGAGCAACCCAGCCCGCGCCACCGACCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyGlyValaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCCCTGCGCGCGCGCCAGAGAGCTGCCACGCCAGGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnLeuProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCTTCAGCCCTGGAGACCAAGAACCTAGGCAAGCAGCCATCTTAGGACA 1113
QY 261 LeuGlnAspAsnLeuLeuGluMetAspIleuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACCAACCTCATGTAGATGGACATCTCGCGCGCTCCCGCCCATGATGGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACITTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCAATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 1234 CCAGGCTACCACTGACGTCAGCACAGTGTACTCTCTCGCGCCGACCCCTGCCCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCCCGCGCGCTTTAACTCAAGAGCCCTCCCAAGTACTGTAACTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCGACCCCTGAGCGCATCGTCATCTCAGCACCTCTGTCATCTCTGGCACTTT 1413
QY 361 ValAlaMetHisIlePheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCATGCACCTGTTGGCTTAACTGGACCTGCAGCGCATGGAGGGCCAGATGTAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCTTGGCTGTGCTTGGCAACCCAGCGTCTCCCTATACCC 1533
RESULT 6
ABN85378
ID ABN85378 standard; DNA; 8438 BP.
AC ABN85378;
XX
XX 21-OCT-2002 (first entry)
DE Human NOV1, TEN-M4 like protein, coding sequence.
KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW TEN-M4 like protein; chromosome 11; gene; ds.
OS Homo sapiens.
FH Key
FT Location/Qualifiers
FT 4. .8395
FT /*tag= a
FT /trans except= (pos: 1138..1147,aa:Met)
FT /product= "NOV1 protein"
XX
XX WO200255704-A2.
XX
XX 18-JUL-2002.
XX
XX 09-JAN-2002; 2002WO-US000554.
```

```
XX 09-JAN-2001; 2001US-0260417P.
PR 10-JAN-2001; 2001US-0260831P.
PR 28-FEB-2001; 2001US-0272338P.
PR 09-MAR-2001; 2001US-0274876P.
PR 18-APR-2001; 2001US-0284704P.
XX (CURA-) CURAGEN CORP.
XX Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Szytek KA,
XX Zhang M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM,
XX Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM,
XX Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K,
XX Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G,
XX Gunther E, Stone DJ;
XX MPI; 2002-590674/63.
XX P-PSDB; ABB98401.
XX NOVX polypeptides and encoding polynucleotides, useful for preventing or
XX treating NOVX-associated disorders e.g. cancer, inflammation, or
XX Alzheimer's disease, and in chromosome mapping, tissue typing or
XX pharmacogenomics.
XX Claim 9; Page 8-9; 358pp; English.
XX The present sequence is a coding sequence for a NOV protein. The NOV
XX proteins and coding sequences are useful for treating or preventing NOV-
XX associated disorders or in the manufacture of a medicament for treating
XX the disorders, such as cancer, heart disease, inflammation, autoimmune
XX disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
XX IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
XX (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
XX and other wasting disorders associated with chronic diseases. NOV1 is a
XX TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
XX
XX Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 5,17e-124 Length: 8438
XX Score: 2082.00 Matches: 395
XX Percent Similarity: 98.01% Conservative: 0
XX Best Local Similarity: 98.01% Mismatches: 5
XX Query Match: 97.65% Indels: 3
XX DB: Gaps: 1
XX
XX SEQ14-X-AT-28-64-76 (1-400) x ABN85378 (1-8438)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 4 ATGGACGTGAGGAGAGGAAGCTTACCGCTCGCTGACCGCGCGCGACGCGGAGCGC 63
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 64 CGCTACACAGCTCGCTCGCGCGAGACAGAGAGGGGCAAGAGCCCGCAGAAATCGTACAGC 123
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 124 TCCAGCGAGACCTTGAAGGCTTACGACAGAGCGCGCGCTAGCCTATGGCAGCCCGCTC 183
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 184 AAGGACATTGTGCGCAGGAGCGCGAGGAATTTCTCGCGCACAGGTGCCAACTTCCACCTG 243
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuThrArgThrAspIle 100
Db 244 CGGAGAGCTGGGGCTGGAAGAGAGTAACGCCCTCTACGGGACCTCTACCGGACAGACAT 303
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 304 GGCCTCCCCCACTGGGGCTACTCCATGGGGCTGGCTCTGATGCGACATGGAGGCTGAC 363
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
```

| | | | | | |
|--|------|--|------|----|--|
| Db | 364 | ACGGTGTCTCCCTGAGCACCCTCGTGGTGTGGGGCGGAGACACGCTGAGGCGC | 423 | KW | renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS; |
| Qy | 141 | SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu | 160 | KW | neurological disorder; Alzheimer disease; Parkinson's disease; asthma; |
| Db | 424 | AGCTCTCTGCTGTCCAGCGGGGCAATTCCTCACTACACCTCAGCAGCAGCAGATGAA | 483 | KW | reproductive disorder; infertility; autoimmune disorder; gout; allergy; |
| Qy | 161 | AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro | 180 | KW | inflammatory disorder; acquired immunodeficiency syndrome; uveitis; |
| Db | 484 | AACACTGAGATGATCATCCGGGGGCGCTGAGAACACACGCGGCTCCGAGCCGCGCG | 543 | KW | autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection; |
| Qy | 181 | ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu | 200 | KW | diabetes mellitus; glomerulonephritis; irritable bowel syndrome; |
| Db | 544 | CGCGGCTCTCGACGCCACACCCCCAACAGACACACGCGGCTCCATTAATCTCCCTG | 603 | KW | multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis; |
| Qy | 201 | AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu | 220 | OS | rheumatoid arthritis. |
| Db | 604 | AACCGGGCAACTTCACCGGAGAGCAACCCAGCCCGGCGCCACGAGCACTCGCTC | 663 | XX | Homo sapiens. |
| Qy | 221 | SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu | 240 | XX | WO200272830-A2. |
| Db | 664 | TCCGAGAGCCCCCTCGCGGGCGGCCAGGAGCTGCCAGCCAGCCAGAGAACTGGGTG | 723 | XX | 19-SEP-2002. |
| Qy | 241 | LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr | 260 | XX | 08-FEB-2002; 2002WO-US003715. |
| Db | 724 | CTCAACAGCAATCCCCCTGAGACACAGAACTAGGCAAGCAGCCATCTCTAGGGACA | 783 | XX | 09-FEB-2001; 2001US-0268111P. |
| Qy | 261 | LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTrp | 280 | XX | 23-FEB-2001; 2001US-0271175P. |
| Db | 784 | TTGAGGCAACCTCATTTGAGATGACATTTCTCGGCGCTCCCGCCATGATGGGCTTAC | 843 | XX | 08-MAR-2001; 2001US-0274503P. |
| Qy | 281 | SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer | 300 | XX | 09-MAR-2001; 2001US-0274552P. |
| Db | 844 | AGTACGGGCACTTCTCTTAAGCTTGAGGACCTCCCGCTCTTCTGACACATCA | 903 | PI | (INCY-) INCYTE GENOMICS INC. |
| Qy | 301 | ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg | 320 | PI | Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR; |
| Db | 904 | CCAGGGTACCCACTGAGCTCCAGCACAGTGATCTCTCTCCGCCCGCCCGCCGCGC | 963 | PI | Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BW, Burford N; |
| Qy | 321 | SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys | 340 | XX | Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK; |
| Db | 964 | AGCACTTTGCGCTGCGCGGCTTTAACTTCAAGAGCCCTCCAGTACTGTAACTGGAG | 1023 | DR | WPI; 2002-723356/78. |
| Qy | 341 | CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTrpPhe | 360 | DR | P-FSDB; ABG97359. |
| Db | 1024 | TCCGAGCCCTGAGCGCATCTGCTATCTCAGCCTCTGCTCATCTCTGTCATCTT | 1083 | XX | New human proteins associated with cell growth, differentiation and |
| Qy | 361 | ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet--- | 379 | XX | death, useful for diagnosing, treating or preventing autoimmune or |
| Db | 1084 | GTGGCCATGCACCTCTTTGGCTTAACTGGCACCTGCAGCCGATGAGGGGCGAGATGAG | 1143 | XX | inflammatory disorders (e.g. AIDS, allergy or anemia), cancer, |
| Qy | 380 | -----TyrGluThrGluAspThrAlaSerSerTrpProValProThrAspValSerIle | 398 | XX | atherosclerosis or hepatitis. |
| Db | 1144 | GATTATGATGATCAGGAGGACACAGCCAGGAGTGGCGCTGTGCCAACCGAGCTCTCCCT | 1203 | PS | Claim 5; Page 175-178; 181pp; English. |
| Qy | 398 | utyrPro 400 | | XX | The invention relates to an isolated polypeptide comprising CGDPI-12 |
| Db | 1204 | ATACCC 1210 | | XX | (cell growth, differentiation and death), a naturally occurring amino |
| RESULT 7 | | | | XX | acid sequence at least 90% identical to CGDP, a biologically active |
| ID ABS78652 | | | | XX | fragment or an immunogenic fragment. Also included are the |
| XX | | | | XX | polynucleotides encoding CGDPI-12, a recombinant polynucleotide |
| AC | | | | XX | comprising a promoter sequence operably linked to the CGDP |
| XX | | | | XX | polynucleotides, a cell transformed with the recombinant polynucleotide, |
| XX | | | | XX | an anti-transgenic organism comprising the recombinant polynucleotide or are |
| XX | | | | XX | CGDP antibody, screening for compounds which bind to/modulate or are |
| XX | | | | XX | antagonists of CGDP or alter the expression of CGDP polynucleotide and a |
| DT 16-DEC-2002 (first entry) | | | | XX | CGDP polynucleotide microarray. The polypeptides, polynucleotides, |
| DE Human cDNA encoding CGDPI10, INCYTE 7488573CB1. | | | | XX | agonists and antagonists are useful for diagnosing, treating or |
| XX | | | | XX | preventing disorders associated with aberrant expression of CGDP, |
| KW | | | | XX | particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis, |
| KW | | | | XX | cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia |
| KW | | | | XX | vera, psoriasis, primary thrombocytopaenia or cancer), developmental |
| KW | | | | XX | disorders (e.g. renal tubular acidosis, anaemia or mental retardation), |
| XX | | | | XX | neurological disorders (e.g. Alzheimer disease, Parkinson's disease or |
| XX | | | | XX | epilepsy), reproductive disorders (e.g. infertility or a disruption in |
| XX | | | | XX | the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS, |
| XX | | | | XX | (acquired immunodeficiency syndrome) allergies, asthma, autoimmune |
| XX | | | | XX | thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus, |
| XX | | | | XX | glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis, |
| XX | | | | XX | osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis, psoriasis, |
| XX | | | | XX | uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic |
| XX | | | | XX | infections. They are also useful in the assessment of the effects of |
| XX | | | | XX | exogenous compounds on the expression of nucleic acid and amino acid |
| XX | | | | XX | sequences of proteins associated with CGDP. The present sequence encodes |
| XX | | | | XX | a CGDP protein |
| XX | | | | XX | Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other; |
| XX | | | | XX | Alignment Scores: 5.67e-123 Length: 8645 |
| XX | | | | XX | Pred. No.: 2066.00 Matches: 390 |
| XX | | | | XX | Score: |

Percent Similarity: 97.50% Conservative: 0
Best Local Similarity: 97.50% Mismatches: 6
Query Match: 96.90% Indels: 4
DB: 6 Gaps: 2

SEQ14-X-AT-28-64-76 (1-400) x ABS78652 (1-8645)

Qy 1 MetAspValLysGluArgLysProThrArgSerLeuThrArgArgAspAlaGluArg 20
Db 117 ATGGACGTGAGGAGAGAGAGCCCTTACCGCTCGCTGAGCCCGCGCGAGCGAGCGC 176

Qy 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 177 CGCTACACCACTCGTCGCGGACAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 236

Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 237 TCCAGCGAGACCCCTGAAGGCTTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 296

Qy 61 LysAspIle***ProGlnGluAlaGluGlyPheCysArgThrGly***AsnPheThrLeu 80
Db 297 AAGGACATTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 356

Qy 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 357 CGGAGGCTGGGCTGGAAGAGTAAACGCCCTCACGGGACCCCTGTACCGGACAGACATT 416

Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 417 GGCCTCCGCCCTCGGCTACTCCATCGGGGCTGGCTCTGATCGGACATGAGGGCTGAC 476

Qy 121 ThrValLeuSerProGluHisProValArgLeuThrProGlyArgSerThrArgSerGlyArg 140
Db 477 ACGGTGCTGTCCCTGAGACCCCGTGGCTGTGGGGGCGGAGCACACCGTTCAGGGCGC 536

Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 537 AGCTCTCGCTGTCAGCGCGGCAATCCCACTCTACACTCACCGACACCGAGCATGAA 596

Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 597 AACACTGAGACT-----CCGGGCGGCTGCAAGAACCCAGCGGGCTTCCGAGCGCGCG 650

Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
Db 651 CGCGCGCTCTCGACGCCACACCCCAACAGACACCGCGGGCTCCATTAACCTCCCTG 710

Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 711 AACCGGGCAACTTCAACCGGAGGAGCAACCCCGAGCGCGGCGGCTTCCGAGCGCGCTC 770

Qy 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnThrLeu 240
Db 771 TCCGAGAGCCCTCGCGCGCGCGCGAGGAGGCTTCCCGACCGCGGAGGAGGAGGAGG 830

Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 831 CTCAACAGCAACATCCCTCTGGAGACAGAACTAGGCAAGCAGGCACTTCTTAGGGACA 890

Qy 261 LeuGlnAspAsnLeuLeuMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 891 TTGACGAGCAACCTATTGAGATGGACATTCTCGGCGCTCCCGCATGATGGGGTTAC 950

Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 951 AGTACGGGCACTTCTCTTCAAGCTGGAGGACCTCCCGCTCTCTGACCAATCA 1010

Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1011 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTTCTTCCCGCGCGAGCCCTTCCCGCG 1070

Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340

Db 1071 AGCACCTTGCCTCGCGCGCGCGCTTTAACTCAAGAGCCCTCCACGACTGTAACTGGAAG 1130

Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360

Db 1131 TGGCAGCGCTTGGCGGCGCATCTGATCTCAGCCACTCTGGTCATCTCTGTCATCTTT 1190

Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

Db 1191 GTGGGTAAACCACTCTTC-----AACTGGCACCTGACGCGATGAGGGGCGAGATGAT 1244

Qy 381 GluLeuThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

Db 1245 GAGATCAGCGAGGACACAGCCAGGCTTGGCTGTGCTGTCACCGAGCTCTCCCTATACCC 1304

RESULT 8
AAV19252
ID AAV19252 standard; cDNA; 2387 BP.
XX AAV19252;
AC AAV19252;
DT 17-AUG-1998 (first entry)
XX Human gamma-hergulin cDNA clone 20.
DE Gamma-hergulin; gamma-HRG; human; autocrine growth factor;
XX breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor;
KW cell proliferation; cell differentiation; cell survival;
KW neurological disorder; muscular disorder; ss.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH CDS 307..1890
FT /*tag= a
FT misc_feature 1186..1263
FT /*tag= b
FT /*note= "insert DNA"
XX WO9802541-A1.
XX 22-JAN-1998.
XX 08-JUL-1997; 97WO-US011841.
XX 12-JUL-1996; 96US-0021640P.
XX (GETH) GENENTECH INC.
XX Schaefer GM, Sliwowski M;
XX WPI; 1998-110589/10.
DR P-PSDB; AAW44818.
XX DNA encoding gamma-hergulin - used to activate ErbB receptor and to
enhance proliferation, differentiation or survival of a cell.
XX Disclosure; Page 56-57; 81pp; English.
XX Clone 20 encodes an N-terminally truncated isoform (see AAW44818) of
human gamma-hergulin (gamma-HRG) (see also AAW44817), a novel member of
the heregulin superfamily. It was isolated from a MDA-MB-175 human breast
cancer cell library after screening with an EGF-like domain. Compared to
gamma-HRG cDNA (see AAV19251), clone 20 contains a 78 base insert between
codons 560 and 561, and lacks the 5' end. The truncated gamma-HRG was
insoluble when expressed as a thioredoxin fusion protein in Escherichia
coli cells. Gamma-HRG nuclear acid and polypeptide are claimed, together
with their uses e.g. for enhancing the proliferation, differentiation or
survival of glial or muscle cells
XX Sequence 2387 BP; 641 A; 640 C; 550 G; 556 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 2.04e-73 Length: 2387

Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.27% Indels: 0
DB: 2 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x AAV19252 (1-2387)

Qy 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCGGCGGCTGTGAGAACACACGCGGGCTCGGAGCGCGCGCGCGCTCTCGCAC 60
Qy 186 AlaHisThrProAsnGlnHisAlaSerLeuAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCCACACCCCAACACGACACACGCGGCTCCATTAACTCCCTGAAACCGGGGCACTTC 120
Qy 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACGCGAGAGGACACCCAGCGCGGCCCCACGAGACCACTCGCTCTCGGAGAGCCCT 180
Qy 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuAsnSerHisPhe 245
Db 181 GCCGCGCGCGCGCGCGCGCTGCCACGCCAGGAGAACTGCTCTCAACAGCAACATC 240
Qy 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAsnLeu 265
Db 241 CCCCTGGAGACAGAAACCTAGGCAAGCAGCCATCTTAGGACATTCGAGGACACCTC 300
Qy 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGATGGACATCTCGCGCGCTCCCGCCATGATGGGCTTACAGTGACGGGCACTTC 360
Qy 286 LeupheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 361 CTTCTTCAAGCTTGGAGGACCTCCCGCTCTTCTGCACCAATCACCAGGGTACCCACATG 420
Qy 306 ThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 ACGTCAGCAGAGTGTACTCTCTCGCGCCGACCCCTCGCGCGGACCTTCGCGCGG 480
Qy 326 ProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLysCysAlaLeuSer 345
Db 481 CCGGCTTTAACTCAAGAGCCCTCAAGTACTGTAACTGGAAGTGGCGAGCCCTGAGC 540
Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365
Db 541 GCCATCGTCATCTCAGCCACTCTGCTCATCTCTGTCGATCTTTGTGGCCATGACCTG 600
Qy 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
Db 601 TTGGGCTTAACTGGGACCTTGCAGCCGATGGAGGGGCGAGTGTATGAGATCAGGAGGAC 660
Qy 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCCAGCAGTGTGGCTGTGGCAACCGACGCTCTCCCTATACGCC 705

RESULT 9

AD807179
ID ADE07179 standard; DNA; 1534 BP.

XX AC ADE07179;

XX DT 29-JAN-2004 (first entry)

XX DE Novel coding sequence (useful for identifying genetic disorders) #245.
XX KW novel gene; novel protein; tissue marker; molecular weight marker;
XX KW chromosome marker; genetic disorder; gene; ds.
XX OS Unidentified.
XX FN WO2003054152-A2.

PD 03-JUL-2003.
XX 10-DEC-2002; 2002WO-US039555.
XX 10-DEC-2001; 2001US-0339739P.
PR 11-DEC-2001; 2001US-0339453P.
PR 14-MAR-2002; 2002US-0365091P.
PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372381P.
PR 12-APR-2002; 2002US-0372615P.
PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Wang G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX WPI; 2003-569235/53.
DR P-PSDB; ADE08090.
XX New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX Claim 1; SEQ ID NO 245; 1177pp; English.
XX The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present DNA sequence represents a gene of the
XX invention.
XX Sequence 1534 BP; 462 A; 377 C; 355 G; 340 T; 0 U; 0 Other;
SQ Alignment Scores:
Pred. No.: 1,04e-53 Length: 1534
Score: 975.00 Matches: 210
Percent Similarity: 62.41% Conservative: 49
Best Local Similarity: 50.60% Mismatches: 92
Query Match: 45.73% Indels: 64
DB: 9 Gaps: 12
SEQ14-X-AT-28-64-76 (1-400) x ADE07179 (1-1534)
Qy 1 MetAspValIleGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 133 ATGGATGTGAAGAACGCGAGCGCTTACTGCTCCCTGACCAAGAGCAGCAGAGAGGAA 192
Qy 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 193 CGGCGCTACACAAATTCCTCCGACAGCAATGAGAGTGGCGGTACCCACAGAGTCC 252
Qy 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 253 TACAGTTCCAGCGAGACATTTAAAGCTTTTGTATCATGATTCCTCGCGGCTGCTTACGGC 312
Qy 58 SerArgValIleAspIle***ProGlnGluAlaGluPheCysArgThrGly***Asn 77
Db 313 AACAGAGTGAAGATTTGGTTTCACAGAGACGAGAGGTTCATCTCGCGGCTGCTTACGGC 372
Qy 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
Db 373 TTTACCTCAGGCGATTTAGGAGTTTGTGAACACGACACTCGAAGAGGACTGGCATTTGT 432
Qy 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 433 GCGGAATGGGCTCCCTCACAGAGGTTACTTATCATGTCAGGGTCAGATGCTGACT 492

CC and protozoal infections. The NOVX proteins can be used as immunogens to
 CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
 CC used in chromosome mapping, identifying individuals from minute
 CC biological samples (tissue typing), and in forensic identification of a
 CC biological sample. The present sequence encodes human NOV15b, which is
 CC located on chromosome 4
 XX
 SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores: 7.52e-53 Length: 8645
 Pred. No.: 975.00 Matches: 210
 Score: 62.41% Conservatives: 49
 Percent Similarity: 50.60% Mismatches: 92
 Best Local Similarity: 50.60% Indels: 64
 Query Match: 6 Gaps: 12
 DB:

SEQ14-X-AT-28-64-76 (1-400) x ABQ82344 (1-8645)

Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
 Db 151 ATGGATGTGAAGAACGAGCGCTTACTGCTCCCTGACCAAGAGCAGACGAGAGGAA 210

Qy 20 ArgArgTyrThrSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
 Db 211 CGGCGCTACACAAATTCCTCCGACACATAGAGGAGTCCGGGTACCCACACAGAAGTCC 270

Qy 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
 Db 271 TACAGTTCACGAGACATGAAAGCTTTGATCATGATTCCTCGCGGTCTTTACGCG 330

Qy 58 SerArgValLysAspLys***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
 Db 331 AACAGAGTGAAGGATTTGGTTCCACAGAGAACGAGAGGAGTTCCTAGACAGACAGAT 390

Qy 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
 Db 391 TTTACCTTAAGCAGTTAGAGTTTGTGAACACGACCACTCGAAGGAGCTGGCATTTTGT 450

Qy 98 ThrAspIleGluProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
 Db 451 GCGGAATGGGCTCCCTCACAGAGTTACTTATCATGTCAGAGGTTCAGATGCTGATCT 510

Qy 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137
 Db 511 GAAATGAAGCAGTGTATCCACAGACATGCGCATGAGACTTTGGGCGCAGGGGTCAAA 570

Qy 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
 Db 571 TCAGCGCGCAGCTCTTCCTGCTGTCAAGTCGGTCCAACTCAGCCCTCACNCTGCAGATACG 630

Qy 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
 Db 631 GAGCAGGAAACAAAGTCCGACAGTGAAGYAGCAACCTGCAAGCAATCAAGCCAGCTCT 690

Qy 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisAla 194
 Db 691 ACCCTGAGCGCTTGGCGCT-----TCCATAAGCAGCAGCTCTGCACAGCATCAT--- 741

Qy 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
 Db 742 CCATCATCACTCTCTCAACAGAACTCCCTGACCAATAGAGGAACACAGAGTCCGGCC 801

Qy 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
 Db 802 CCG-----CCGCGCTGCTTTGCCCGCGAGCTGCAACCAACCA 837

Qy 231 ---GluProAlaHisAlaGlnGluAsnThrLeuAsnSerAsnIleProLeuGluThr 249
 Db 838 CCGGAGTCCGTCAGCTGCAGCAGCTGGGTCTTGGCAGTAAATGACCATGGAAAGC 897

Qy 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
 Db

898 AGG----- 900
 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
 901 -----CATTTCTATTCAAAACA 918
 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
 919 GGAACAGGTACACGCCACTGTTTCAGTACTGCAACCCAGGATACACAATGGCATCTGGC 978
 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
 979 TCTGTTTATTCACCACTTCTCGGCCACTACTAGAACACCCCTATCAAGAGTCTTT 1038
 329 AsnLeuLysLysProSerLysTyrCysAsnTyrLysCysAlaAlaLeuSerAlaIleVal 348
 1039 AAATTCAGAAAGTCTTCAAGTACTGTAGCTGGAATGCACTGCACCTGTGCGGTAGGG 1098
 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
 1099 GTCTCGGTGCTCTGGCAATATCTCTGCTTATTATTATAGCAATGCATCTCTTTGGCCTC 1158
 369 AsnTyrHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
 1159 AACTGGCAGCTACAGCAGACTGAAATGACACATTTGAGATGGAAGTGAATTCGTAT 1218
 386 ThrAlaSerSerTyrProValProThrAspValSerLeuTyrPro 400
 1219 ACC-----ATGCCAACAAACACTGTGTCTATTACCT 1248

RESULT 11
 ABQ82343
 ID ABQ82343 standard; cDNA; 8675 BP.
 XX
 AC ABQ82343;
 XX
 DT 17-DEC-2002 (first entry)
 XX
 DE Human NOV15a encoding cDNA SEQ ID NO:35.
 XX
 KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
 KW cerebroprotective; nontropic; antidiabetic; antiinflammatory; fungicide;
 KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
 KW antianemic; antibacterial; protozoacide; antihelminthic; gene therapy;
 KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
 KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
 KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
 KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
 KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
 KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
 KW rheumatoid arthritis; gene; chromosome 4; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 151..8328
 FT /*tag= a
 FT /product= "NOV15a"
 FT /transl_except= (pos:1249..1251,aa:Ser)
 XX
 PN WO200262999-A2.
 XX
 PD 15-AUG-2002.
 XX
 PF 31-DEC-2001; 2001WO-US049976.
 XX
 PR 29-DEC-2000; 2000US-0258928P.
 PR 02-JAN-2001; 2001US-0259415P.
 PR 04-JAN-2001; 2001US-0259785P.
 PR 20-FEB-2001; 2001US-0269814P.
 PR 09-MAR-2001; 2001US-0279863P.
 PR 29-MAR-2001; 2001US-0279832P.
 PR 29-MAR-2001; 2001US-0279833P.

13-APR-2001; 2001US-0283889P.
 PR 18-APR-2001; 2001US-0284447P.
 PR 25-APR-2001; 2001US-0286683P.
 PR 29-MAY-2001; 2001US-0294080P.
 PR 16-AUG-2001; 2001US-0312915P.
 PR 17-AUG-2001; 2001US-0313325P.
 PR 17-SEP-2001; 2001US-0322699P.
 PR 26-NOV-2001; 2001US-0333350P.
 XX (CURA-) CURAGEN CORP.
 XX
 XX Szytek KA, Li L, Wolenc AR, Vernet CM, Eisen A, Liu X;
 PI Malyskar U, Shimkova RA, Tchernov VT, Spaderna SK, Gorman L;
 PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
 PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Biding S, Ellerman K;
 PI Gunther E, Smithson G, Millet I, Macdougall JR;
 XX
 XX WFI; 2002-732706/79.
 DR P-PSDB; ABP53586.
 XX
 XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
 PT associated disorders, such as cancers, neurological disorders, disorders
 PT of vesicular transport, gastrointestinal disorders, and autoimmune
 PT diseases.
 XX
 XX Claim 8; Page 110-112; 444pp; English.
 XX
 XX The present invention describes novel human proteins designated NOVX,
 CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
 CC cytotatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
 CC antiinflammatory, antirheumatic, antithrombotic, virucide,
 CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
 CC protozoacide and antihelminthic activities, and can be used in gene
 CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease selected from NOVX-associated disorder, such as cancers
 CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
 CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
 CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
 CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
 CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
 CC disorders (e.g. ulcerative colitis, or gastritis), autoimmune disorders
 CC (autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
 CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
 CC and protozoal infections). The NOVX proteins can be used as immunogens to
 CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
 CC used in chromosome mapping, identifying individuals from minute
 CC biological samples (tissue typing), and in forensic identification of a
 CC biological sample. The present sequence encodes human NOV15a, which is
 CC located on chromosome 4
 XX
 SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;

Alignment Scores:

| Pred. No.: | 7.55e-53 | Length: | 8675 |
|------------------------|----------|---------------|------|
| Score: | 975.00 | Matches: | 210 |
| Percent Similarity: | 62.41% | Conservative: | 49 |
| Best Local Similarity: | 50.60% | Mismatches: | 92 |
| Query Match: | 45.73% | Indels: | 64 |
| DB: | 6 | Gaps: | 12 |

SEQ14-X-AT-28-64-76 (1-400) x ABQ82343 (1-8675)

QY 1 MetAspValLysGluArgLysProTyrArgSerGluThrArg---ArgArgAspAlaGlu 19
 DB 151 ATGGATGTGAAGAACGCGAGGCTTACTGCTCCCTGACCAAGAGCAGACAGAGAAGGAA 210
 QY 20 ArgArgTyrThrSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
 DB 211 CGGCGCTACACAATCTCCGACACATATGAGGAGTCCGGGTACCCACACAGAGTCC 270
 QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
 DB 151 ATGGATGTGAAGAACGCGAGGCTTACTGCTCCCTGACCAAGAGCAGACAGAGAAGGAA 210

| Db | QY | 271 | TACAGTTCACGCGACGACATTGAAAGCTTTTGATCATGATCTCTCGCGGCTGCTTACGCGC 330 |
|----|------|---|---|
| QY | 58 | SerArgValLysAspIle***ProGlnGluAlaGluPheCysArgThrGly***Asn 77 | 58 SerArgValLysAspIle***ProGlnGluAlaGluPheCysArgThrGly***Asn 77 |
| Db | 331 | AACAGAGTGAAGATTGGTTCACAGAGACAGACAGAGTTCACAGACAGAGTTCACAGACAGAT 390 | 331 AACAGAGTGAAGATTGGTTCACAGAGACAGACAGAGTTCACAGACAGAGTTCACAGACAGAT 390 |
| QY | 78 | PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97 | 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97 |
| Db | 391 | TTTACCTTAGGCGAGTTTGTGAACCACTCGAAGAGGAGTGGCATTTGT 450 | 391 TTTACCTTAGGCGAGTTTGTGAACCACTCGAAGAGGAGTGGCATTTGT 450 |
| QY | 98 | ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117 | 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117 |
| Db | 451 | GCGAAATATGGGCTCCCTCAGAGAGTACTCTATCATGTGCGAGGTTCAGATGCTGATACT 510 | 451 GCGAAATATGGGCTCCCTCAGAGAGTACTCTATCATGTGCGAGGTTCAGATGCTGATACT 510 |
| QY | 118 | GluAlaSerThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137 | 118 GluAlaSerThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137 |
| Db | 511 | GAAATGAACGAGTGTGCTCCAGAGATGCTCCAGAGATGCTCCAGAGATGCTCCAGAGATGCT 570 | 511 GAAATGAACGAGTGTGCTCCAGAGATGCTCCAGAGATGCTCCAGAGATGCTCCAGAGATGCT 570 |
| QY | 138 | SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157 | 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157 |
| Db | 571 | TCAGGCGCAGCTCTGCTCAAGTCGGTCCAACTCAGCCCTCAGCCCTCAGCCCTCAGAGATACG 630 | 571 TCAGGCGCAGCTCTGCTCAAGTCGGTCCAACTCAGCCCTCAGCCCTCAGCCCTCAGAGATACG 630 |
| QY | 158 | GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174 | 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174 |
| Db | 631 | GAGCAGAAACAAAGTCCGACGAGTGAAGATGACCACTGACCACTGACCACTGACCACTGAC 690 | 631 GAGCAGAAACAAAGTCCGACGAGTGAAGATGACCACTGACCACTGACCACTGACCACTGAC 690 |
| QY | 175 | ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194 | 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194 |
| Db | 691 | ACCTGCGAGCTTGGCGCT-----TCCATAGCAGCAGCAGCTGTCACAGCATCAT--- 741 | 691 ACCTGCGAGCTTGGCGCT-----TCCATAGCAGCAGCAGCTGTCACAGCATCAT--- 741 |
| QY | 195 | AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214 | 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214 |
| Db | 742 | CCATCCATCACTTCTCTCAACAGAACTCCCTGACCAATAGAGAACACAGAGTCCGCC 801 | 742 CCATCCATCACTTCTCTCAACAGAACTCCCTGACCAATAGAGAACACAGAGTCCGCC 801 |
| QY | 215 | ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230 | 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230 |
| Db | 802 | CCG-----CCGGCTGCTTTCGCCCGCAGCTGCAAAACCA 837 | 802 CCG-----CCGGCTGCTTTCGCCCGCAGCTGCAAAACCA 837 |
| QY | 231 | ---GluProAlaHisAlaGlnGluAsnTyrLeuLeuAsnSerAsnIleProLeuGluThr 249 | 231 ---GluProAlaHisAlaGlnGluAsnTyrLeuLeuAsnSerAsnIleProLeuGluThr 249 |
| Db | 838 | CCGAGTCCGCTCCAGCTGAGGACAGCTGGTCTTGGCAGATGATGATACCTGGAAGAC 897 | 838 CCGAGTCCGCTCCAGCTGAGGACAGCTGGTCTTGGCAGATGATGATACCTGGAAGAC 897 |
| QY | 250 | ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuLeuGluMetAsp 269 | 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuLeuGluMetAsp 269 |
| Db | 898 | AGG----- 900 | 898 AGG----- 900 |
| QY | 270 | IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289 | 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289 |
| Db | 901 | -----CATTCCTATTCAAAACA 918 | 901 -----CATTCCTATTCAAAACA 918 |
| QY | 290 | Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308 | 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308 |
| Db | 919 | GGAACAGGTACAAACGCCACTGTTAGTACTGCAACCCAGGATACAAATGATGCTGCG 978 | 919 GGAACAGGTACAAACGCCACTGTTAGTACTGCAACCCAGGATACAAATGATGCTGCG 978 |
| QY | 309 | ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328 | 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328 |
| Db | 979 | TCTGTTTATTCACCACTACTCGGCCACTACCTAGAAACACCCCTATCAAGAAAGTCTTT 1038 | 979 TCTGTTTATTCACCACTACTCGGCCACTACCTAGAAACACCCCTATCAAGAAAGTCTTT 1038 |
| QY | 329 | AsnLeuLysLysProSerLysTyrCysAsnTyrLysCysAlaAlaLeuSerAlaVal 348 | 329 AsnLeuLysLysProSerLysTyrCysAsnTyrLysCysAlaAlaLeuSerAlaVal 348 |
| Db | 1039 | AAATTCAGAGAGTCTTCAAGTACTGTAGTGGAAATGCATGCACTGTGTCGCTAGGG 1098 | 1039 AAATTCAGAGAGTCTTCAAGTACTGTAGTGGAAATGCATGCACTGTGTCGCTAGGG 1098 |
| QY | 349 | IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368 | 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368 |
| Db | 1099 | GTCTCGGCTCTCTGGCAATACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1158 | 1099 GTCTCGGCTCTCTGGCAATACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1158 |
| QY | 369 | AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385 | 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385 |
| Db | 1159 | AACTGGCAGCTACAGCAGACTGAAATGACACATTTGAGATGGAAGAAAGTGAATTCAT 1218 | 1159 AACTGGCAGCTACAGCAGACTGAAATGACACATTTGAGATGGAAGAAAGTGAATTCAT 1218 |
| QY | 386 | ThrAlaSerSerTyrProValProThrAspValSerLeuTyrPro 400 | 386 ThrAlaSerSerTyrProValProThrAspValSerLeuTyrPro 400 |
| Db | 1219 | ACC-----ATGCCAACAAACACTGTGTCAATACCT 1248 | 1219 ACC-----ATGCCAACAAACACTGTGTCAATACCT 1248 |

RESULT 12
 ACCT72052
 ID ACCT72052 standard; DNA; 9695 BP.
 XX
 AC ACCT72052;
 XX
 DT 08-JUL-2003 (first entry)
 XX
 DE BC0205B gene #SEQ ID 81.
 XX
 KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
 KW drug discovery; clinical medicine; forensic medicine; gene;
 KW chromosome 5q33.3; ds.
 XX
 XX Homo sapiens.
 XX
 XX W02003029421-A2.
 XX
 XX 10-APR-2003.
 XX
 XX 02-OCT-2002; 2002WO-US031287.
 XX
 XX 03-OCT-2001; 2001US-0326526P.
 PR 14-MAY-2002; 2002US-00144194.
 XX
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX PA
 XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;
 PI
 XX WPI; 2003-381623/36.
 DR P-PSDB; ABR58318.
 XX
 XX New isolated human differentially-regulated breast cancer polynucleotide
 PT and polypeptide, useful for diagnosing, staging, prognosticating,
 PT preventing and/or treating diseases and conditions relating to breast
 PT cancer.
 XX
 XX Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.
 PS
 CC The invention relates to isolated polynucleotides which are
 CC differentially-regulated in breast cancer. The methods and compositions
 CC of the present invention are useful for detecting, diagnosing, staging,
 CC monitoring, prognosticating, preventing and/or treating diseases and
 CC conditions relating to breast cancer, and may be used in gene therapy or
 CC antisense therapy. They can also be used in research, drug discovery,
 CC clinical medicine and forensic medicine. Sequences given in records
 CC ACC72012-ACC72074 represent polynucleotides of the invention that are
 CC differentially-regulated in breast cancer. NOTE: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 2,38e-50 Length: 9695
 Score: 937.00 Matches: 206
 Percent Similarity: 63.57% Conservative: 54
 Best Local Similarity: 50.37% Mismatches: 93
 Query Match: 43.95% Indels: 56
 DB: 7 Gaps: 13

SEQ14-X-AT-28-64-76 (1-400) x ACCT72052 (1-9695)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArg---AspAlaGlu 19
 Db 435 AUGAGTAAAGACCGGGA---CACCCTTTTGACAGAGACGCTGTGGCAAGAG 491
 QY 20 ArgArgTyrThrSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
 Db 492 TGTGCTACTACAAAGCTCCTCTCTGGACAGTGAAGTGGCGGCTGGCCACAGAAATCC 551

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
 Db 552 TACAGCTCCAGTGAGACTCTGAAGCCCTATGACATGACAGCAGGATGCAATCGAAAC 611
 QY 59 ArgValLysAspLys***ProGlnGluAlaGluPheCysArgThrGly**AsnPhe 78
 Db 612 CGAGTCAAGACCTCATCCACCGGAGTCAAGTGTCTCTAGACAGGACCACTTC 671
 QY 79 ThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThr 98
 Db 672 ACCCTTGGCGACATGGGCACTGTGAGCCCTCC---CCACACCGAAGCGGTACTGTCTCC 728
 QY 99 AspLeuGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
 Db 729 GACATGGGATCCTTCCACAGGCTACTCCCTTAGCACAGGCTCTGACCGCGACTCCGAC 788
 QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSer 138
 Db 789 ACCGAGGAGGAGATGCTCTCCAGAACACGCCATCAGACTGTGGGCGAGAGGATAAATCC 848
 QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
 Db 849 AGGCGCAGTTCGGCCTGTCTCAGTCTGTAACCTCGGCCCTTACCTGACTGACTCTGAC 908
 QY 159 HisGluAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArg-----Leu 176
 Db 909 AACGAAACAAATCAGATGATGAGAACCGCCCTCGAAACACCCACAGCAGTCTGACTCTG 968
 QY 177 ArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSer 196
 Db 969 AGGCCCTCTCCACCC---CCTCACAAACACACGCTGTCCCATCACCAC---TCGTCC 1022
 QY 197 IleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn-----ProSer 212
 Db 1023 GCCAACTCCCTCAACAGGACTCAGTACCAATCGCGAGTGCAGATCCACGCCCGGCC 1082
 QY 213 ProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluPro 232
 Db 1083 CCAGCGCCCAATGAC-----CTGGCCACACACCA-----GAGTCC 1118
 QY 233 AlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeu 252
 Db 1119 GTTCAGCTTCAGACAGCTGGGTGTAAACACAGCTGCCACTGGAGACCCGG-----1172
 QY 253 GlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGly 272
 Db 1172 -----1172
 QY 273 AlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLys---ProGlyGly 291
 Db 1173 -----CACTTCTCTTCAAGACCTCTCGGG 1199
 QY 292 ThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyr 311
 Db 1200 AGCACACCTTGTTCAGCAGCTCTTCCCGGGATACCTTTGACCTCAGAACGGTTTAC 1259
 QY 312 SerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeuLys 331
 Db 1260 ACGCCCGCCCGCCCTGCTGCCAGGAATACTTTCTCCAGGAAGGCTTTCAGAGCTGAAG 1319
 QY 332 LysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSerAla 351
 Db 1320 AAGCCCTCCAAATACTGCAGCTGGAATGTGCGCCTCTCCGCCATTCGCCGCGGCCCTC 1379
 QY 352 ThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrpHis 371
 Db 1380 CTCTTGCTGATTATTGCTGGGTAATTCATAGCAATGATCATCTGCTCGGACTCAATGGCAA 1439
 QY 372 LeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSerSerTrpPro 391
 Db 1440 CTCAGCCTGCAGATGGGCACACCTTT-----AACAAATGGGATAAGGACCGGTTACCA 1493
 QY 392 ValProThrAspValSerLeuTyrPro 400

```

Db      1494 GGAACGATGATGGCAACATGCCA 1520
RESULT 13
AAS14089
ID      AAS14089 standard; DNA; 9729 BP.
AC      AAS14089;
XX
XX      18-DEC-2001 (first entry)
XX
XX      Human FCTR3f DNA sequence.
XX
KW      Human; FCTR3f; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
KW      astrocytoma; congenital neonatal autoimmune thrombocytopenia; infection;
KW      neurological disorder; neurodegenerative disorders; nerve trauma;
KW      familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW      demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW      mental health condition; immunological disorder; allergy; infertility;
KW      bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW      reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW      desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW      gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW      Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW      Corneal dystrophy-Greenow type I; Corneal dystrophy-lattice type I;
KW      Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW      anti-allergic; antiasthmatic; anti-infectivity; anti-inflammatory;
KW      antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
KW      gynaecological; anti-infectivity; immunostimulant; auditory; haemostatic;
KW      gene therapy; FCTR3a; neurastin-like protein; FCTR3f.
XX
OS      Homo sapiens.
XX
XX      Key      Location/Qualifiers
FH      5'UTR      1..209
FT      /*tag= b
FT      CDS        210..8384
FT      /*tag= a
FT      /product= "Human FCTR3f"
FT      3'UTR      8385..9729
FT      /*tag= c
XX
XX      WO200166747-A2.
XX
XX      13-SEP-2001.
XX
XX      05-MAR-2001; 2001WO-US007160.
XX
XX      03-MAR-2000; 2000US-0186592P.
XX      03-MAR-2000; 2000US-0186718P.
XX      06-MAR-2000; 2000US-0187293P.
XX      17-MAR-2000; 2000US-0187294P.
XX      17-MAR-2000; 2000US-0190400P.
XX      07-APR-2000; 2000US-0196018P.
XX      03-JAN-2001; 2001US-0259548P.
XX
XX      (CURA-) CURAGEN CORP.
XX
XX      Vernet CAM, Fernandes E, Shinkets RA, Herrmann JL, Majumder K;
XX      Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
XX      WPI; 2001-596837/67.
XX      P-PSDB; AAU08681.
XX
XX      Novel polypeptides designated as FCTR3 polypeptides, useful in detection,
XX      prevention and treatment of a broad range of pathological states.
XX
XX      Claim 9; Page 37-39; 215pp; English.
XX
XX      The invention relates to human FCTR3 polypeptides, FCTR1-FCTR7, and the
XX      nucleic acids encoding them. These sequences are useful for the treatment
XX      or prevention of numerous disorders including myelogenous leukaemia,
XX      carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal

```

```

CC      all immune thrombocytopenia, neurological disorders, neurodegenerative
CC      disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
CC      -Tooth neuropathy, demyelinating Gardner syndrome, familial
CC      myelodysplastic syndrome, mental health conditions, immunological
CC      disorders, allergy and infection, bronchial asthma, Avellino type
CC      eosinophilia, lung diseases, reproductive disorders, infertility, male
CC      and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC      desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC      disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC      infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
CC      Corneal dystrophy-Greenow type I, Corneal dystrophy-lattice type I and
CC      Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
CC      the FCTR3a homologue FCTR3f
XX
SQ      Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

```

Alignment Scores:

```

Pred. No.:      2,42e-43      Length:      9729
Score:          828.00      Matches:      194
Percent Similarity: 53.98%      Conservative: 50
Best Local Similarity: 42.92%      Mismatches: 78
Query Match:      38.84%      Indels:      130
DB:              5      Gaps:      14

```

SEQ14-X-AT-28-64-76 (1-400) x AAS14089 (1-9729)

```

QY      1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArg---AspAlaGlu 19
DB      210 ATGGAUTAAAGACCGCGCA---CACCGCTCTTTGACAGAGACGCTGTGGCAAGAG 266
QY      20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
DB      267 TGTGCGCTACACAAGCTCTCTCTGGACAGTGAGAGACTCGCGGTGCCACACAGAAATCC 326
QY      39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
DB      327 TACAGCTCCAGTGAGACTCTGAAGGGCTATGACCATGACAGCAGGATGCTATGGAAC 386
QY      59 ArgValLysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPhe 78
DB      387 CGAGTCACAGACTCATCCACGGGAGTCAGATGAGTTTCCTAGACAAAGAACCACTTC 446
QY      79 ThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThr 98
DB      447 ACCCTTGGCGAACTGGGCATCTGTGAGCCTCC---CCACACGAGCGGCTACTGCTCC 503
QY      99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
DB      504 GACATGGGGATCTTCCACAGGGCTACTCCCTTAGCACAGGGTCTGACGCGGACTCCGAC 563
QY      119 AlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSer 138
DB      564 ACCGAGGGAGGATGTCTCCAGAACACGCCATCAGACTGTGGGGCAGAGGATAAATCC 623
QY      139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
DB      624 AGGCGCAGTTCGGGCTGTCCAGTCGTGAAAACTCGGCCCTTACCTGACTGACTGCTGAC 683
QY      159 HisGluAsnThrGluThrAspHisProGly-----
DB      684 AACGAAACAAATCAGATGATGAGAACGGTCGTCCCATTCACCTACATCTCCTCGCTAGT 743
QY      168 -----
DB      744 CTCCTCCCATCTGCTCAGCTCCTAGTCCCATATCTCCACCGATTAGTGTCCAGATG 803
QY      168 -----
DB      804 CCATTGCTAGACAGCAACACCTCCCATCAAATCATGGACACCAACCCCTGATGAGGAATTC 863
QY      169 -----
DB      864 TCCCCCAATTCACTGCTCAGAGCATGCTCAGGGCCCCAGCAAGCCTCCAGCAGTGGC 923

```

QY 170 LeuGluAsnHis-----AlaArgLeuArgThrProProProProLeuSerHisAla 186
 Db 924 CCTCGGAACACACACAGCCAGTCTGAGGCGCCCTCTCCACCC---CCTCACAAC 980
 QY 187 HisThrProAsnGlnHisHisAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
 Db 981 CACACGCTGCCATCACAC---TCGTCCGCCAACTCCCTCAACAGAACTCACTGACC 1037
 QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
 Db 1038 AATCGCGGAGTCAGATCCAGCCCGCCGCCCGCCCAATGAC-----CTGCCACAC 1091
 QY 223 GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsn 242
 Db 1092 ACACCA-----GAGTCCCTTCAGCTTCAGGACAGCTGGGTGCTAAAC 1133
 QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyGlnProPheLeuGlyThrLeuGln 262
 Db 1134 AGCAACGTGCCACTGGAGACCCGG----- 1157
 QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTySerAsp 282
 Db 1157 ----- 1157
 QY 283 GlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
 Db 1158 ---CACTTCTCTTCAAGACCTCTCTCGGGGAGCACACCTTGTTCAGCAGCTCTTCCCG 1214
 QY 302 GlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArgSer 321
 Db 1215 GGATACCTTGTGACCTCAGGACGGTTTACAGCCCGCCCGCCCTGCTGCCAGGAAT 1274
 QY 322 ThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLysCys 341
 Db 1275 ACTTCTCCAGAGGCTTCAAGCTGAAGAGCCCTCCAAATCTGCAGCTGGAATGT 1334
 QY 342 AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyPheVal 361
 Db 1335 GCTGCCCTCTCGCCATTCCCGCGCCCTCTTGGCTGATTTTGTGCGGATTTTCATA 1394
 QY 362 AlaMetHisLeuPheGlyLeuAsnTrpHisLeuGln 373
 Db 1395 -----GTGCCCTGTGCTTGAAA 1412
 RESULT 14
 ID ADB32028 standard; cDNA; 9729 BP.
 XX AC ADB32028;
 XX 04-DEC-2003 (first entry)
 XX Human FCTR3f cDNA.
 KW Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
 KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
 KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
 KW renal cell carcinoma; melanoma; clear cell carcinoma;
 KW granular cell carcinoma; neurological disorder;
 KW neurodegenerative disorder; nerve trauma;
 KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
 KW Gardner syndrome; mental health condition; immunological disorder;
 KW allergy; asthma; lung disease; reproductive disorder; deafness;
 KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
 KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
 KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
 KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
 KW lattice corneal dystrophy.
 XX OS Homo sapiens.
 XX US2003087816-A1.

XX 08-MAY-2003.
 XX 05-MAR-2001; 2001US-00800198.
 XX 03-MAR-2000; 2000US-0186592P.
 XX (VERM//) VERMET C.
 XX (FERN//) FERNANDES E.
 XX (SHIM//) SHIMKETS R.
 XX (HERR//) HERRMANN J.
 XX (MAJU//) MAJUMDER K.
 XX (MACD//) MACDOUGALL J.
 XX (MISH//) MISHRA V.
 XX (MEZE//) MEZES P S.
 XX (RAST//) RASTELLI L.
 XX Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
 XX Macdougall J, Mishra V, Mezes PS, Rastelli L;
 XX WPI; 2003-625633/59.
 XX P-PSDB; ADB32029.
 XX New FCTRX polypeptide and encoding polynucleotide, useful for preventing
 XX or treating FCTRX-related disorders, such as cancer, autoimmune,
 XX neurodegenerative, gastrointestinal, reproductive and inflammatory
 XX diseases.
 XX Claim 9; Page 34-37; 155pp; English.
 XX The invention relates to FCTRX polypeptides and the polynucleotides
 XX encoding them. The sequences of the invention are useful for the
 XX manufacture of a medicament for diagnosing and treating disorders
 XX associated with the FCTRX polypeptide, such as colorectal cancer,
 XX adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
 XX autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
 XX tumours, mammary tumours, human gliomas, astrocytomas, renal cell
 XX carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
 XX cell and granular cell carcinomas, neuroendocrine disorders,
 XX neurodegenerative disorders, nerve trauma, familial myelodysplastic
 XX syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
 XX conditions, immunological disorders, allergy and infection, asthma, lung
 XX diseases, male and female reproductive disorders, deafness, glycoprotein
 XX deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
 XX C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
 XX infection, spinocerebellar ataxia, plasmodium falciparum infection,
 XX Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
 XX represents cDNA encoding an FCTRX polypeptide of the invention.
 XX Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 2,42e-43 Length: 9729
 Score: 828.00 Matches: 194
 Percent Similarity: 53.98% Conservative: 50
 Best Local Similarity: 42.92% Mismatches: 78
 Query Match: 38.84% Indels: 130
 DB: 9 Gaps: 14
 SEQ14-X-AT-28-64-76 (1-400) x ADB32028 (1-9729)
 QY 1 MetAspValIysGluArgLysProTyrArgSerLeuThrArgArg---AspAlaGlu 19
 Db 210 ATGGATGTAAAGGACCGGGA---CACCGCTCTTTCACAGAGACGCTGTGCAAGAG 266
 QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
 Db 267 TGTGCTACACAGCTCTCTCTGGACAGTGGAGCTGCCGGGTGCCACACAGAAATCC 326
 QY 39 TyrSerSerSerGluThrLeuLysAlaTyAspGlnAspAlaArgLeuAlaTyGlySer 58
 Db 327 TACAGCTCCAGTGAAGCTCTGAAGGCTATGACCATGACAGCAGATGCATATGGAAC 386

CC The invention relates to human FCTR polypeptides, FCTR1-FCTR7, and the
 CC nucleic acids encoding them. These sequences are useful for the treatment
 CC or prevention of numerous disorders including myelogenous leukaemia,
 CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
 CC allolimmune thrombocytopaenia, neurological disorders, neurodegenerative
 CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
 CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
 CC myelodysplastic syndrome, mental health conditions, immunological
 CC disorders, allergy and infection, bronchial asthma, Avellino type
 CC eosinophilia, lung diseases, reproductive disorders, infertility, male
 CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
 CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
 CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
 CC infection Spino cerebellar ataxia, Plasmodium falciparum parasitaemia,
 CC Corneal dystrophy-Greenow type I, Corneal dystrophy-lattice type I and
 CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
 CC FCTR3b, a neurestin-like protein
 XX

SQ Sequence 9826 BP; 2567 A; 2643 C; 2479 G; 2137 T; 0 U; 0 Other;

Alignment Scores: 2.45e-43 Length: 9826
 Pred. No.: 828.00 Matches: 194
 Score: 53.98% Conservative: 50
 Percent Similarity: 42.92% Mismatches: 78
 Best Local Similarity: 38.84% Indels: 130
 Query Match: 5 Gaps: 14
 DB:

SEQ14-X-AT-28-64-76 (1-400) x AAS14085 (1-9826)

QY 1 MetAspVallyGluArgLysProTyrArgSerLeuThrArgArg---AspAlaGlu 19
 DB 280 ATGGATGTAAAGGACCGGCGA---CACCGCTCTTTGACCAGAGGACGCTGTGGCAAGAG 336
 QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGluLysAlaPro---GlnLysSer 38
 DB 337 TGTGCTACACAAAGTCTCTCTGTGACAGTGTGAGGACTGCGGGTGTCCACACAGAAATCC 396
 QY 39 TyrSerSerSerGluThrLysAlaTyrArgGlnAspAlaArgLeuAlaTyrGlySer 58
 DB 397 TACAGCTCCAGTGAGACTGTGAAGCCCTATGACCATGACAGCGAGTGTGACATGTGGAAC 456
 QY 59 ArgVallyAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPhe 78
 DB 457 CGAGTCACAGACTCATCCCGGGAGTCAGATGATTTCTTAGACAAAGAACCAATTC 516
 QY 79 ThrLeuArgGluLeuGluGluValThrProProHisGlyThrLeuTyrArgThr 98
 DB 517 ACCCTTGCGGAACGTGGCATCTGTGAGCCCTCC---CCACACCGAAGCGGCTACTGCTCC 573
 QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
 DB 574 GACATGGGATCTCTCACAGGGCTACTCCCTTAGCACAGGGTCTGACCGCGACTCCGAC 633
 QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSer 138
 DB 634 ACCGAGGGAGGATCTCTCCACACACCGCATCAGACTGTGGGGCAGAGGGATAAAATCC 693
 QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
 DB 694 AGCGCAGTTCGGGCTCTGTCAGTCTGTGAAATCTGGCCCTTACCTGTGACTGACTGAC 753
 QY 159 HisGluAsnThrGluThrAspHisProGly----- 168
 DB 754 AACGAAACAAATCAGATGATGAGAACGGTCTGTCCTCCATTCACCTACATCTCTCGCTAGT 813
 QY 168 ----- 168
 DB 814 CTCCTCCCATCTGCTCAGCTGCTAGTCTCCCAATATCTCCACAGTTAGTGTCCAGATG 873
 QY 168 ----- 168
 DB 874 CCATTGCTAGACAGCAACACCTCCCATCAATCATGAGACCAACCTGATGAGGAATTC 933

Search completed: August 14, 2004, 19:26:05

Job time : 602 secs

QY 169 -----Gly 169
 DB 934 TCCCCCAATTACATCTGCTCAGAGCATGCTCAGGGCCCCCAGCAAGCTCCAGCAGTGGC 993
 QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrProProProProLeuSerHisAla 186
 DB 994 CCTCGGAACCAACACAGCCAGTCAGTCTGAGGGCCCTCTCCACACC---CCTCACAAAC 1050
 QY 187 HisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
 DB 1051 CACACGGTGTCCCATCACCAC---TCGTCCGCCCAACTCCCTCAACAGGAATCTACTGACC 1107
 QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
 DB 1108 AATCGCGGAGTCCAGTCCACGCCCGCCGCGCCAGCGCCCAATGAC-----CTGCCACAC 1161
 QY 223 GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnThrLeuLeuAsn 242
 DB 1162 ACACCA-----GAGTCCGTTTCAGCTTCAGGACAGCTGGTGTCTAAAC 1203
 QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln 262
 DB 1204 AGCAACGTGGCCACTGGAGACCCGG----- 1227
 QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
 DB 1227 ----- 1227
 QY 283 GlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
 DB 1228 ---CACTTCTCTTCAAGACCTCTCGGGGAGCACACCTTTGTCAGCAGCTCTTCCCGG 1284
 QY 302 GlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArgSer 321
 DB 1285 GGATACCTTTGACCTCAGGAACGGTTTACAGGCCCGCCCGCGCTGCTGCCAGGAAT 1344
 QY 322 ThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCys 341
 DB 1345 ACTTTCTCCAGGAAGGCTTTCAAGCTGAAGAACCCCTCCAAATACTGCAGTGTGAAATGT 1404
 QY 342 AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal 361
 DB 1405 GCTGCCCTCTCGCCATTTGCGCGGCCCTCTCTTGGCTATTTTGTGCGGTATTTCATA 1464
 QY 362 AlaMetHisLeuPheGlyLeuAsnTrpHisLeuGln 373
 DB 1465 -----GTGCCCTGTCGTGTGAAA 1482

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 120.435 Seconds

(without alignments)

2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_450_520

Perfect score: 372

Sequence: 1 TFWRSQVFIDHPVHLKFNVS.....RLITQBARSLGFTPROSRGT 71

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10029020/runat_06082004_112215_29265/app.query.fasta_1.13519
-DB=N_Geneseq_29Jan04 -QWTF=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020.gcgn.1.1868 @runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPHLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04.*

1: geneseq1980s.*
2: geneseq1990s.*
3: geneseq2000s.*
4: geneseq2001as.*
5: geneseq2001bs.*
6: geneseq2002s.*
7: geneseq2003as.*
8: geneseq2003bs.*
9: geneseq2003cs.*
10: geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|------------|--------------------|
| 1 | 372 | 100.0 | 2387 | 2 AAV19252 | AAV19252 Human gam |
| 2 | 372 | 100.0 | 3111 | 2 AAV19251 | AAV19251 Human gam |
| 3 | 372 | 100.0 | 3111 | 2 AAX87705 | AAX87705 Gamma-her |
| 4 | 372 | 100.0 | 3111 | 3 AAD00791 | AAD00791 Human Her |
| 5 | 372 | 100.0 | 3111 | 4 AAS18526 | AAS18526 DNA encod |
| 6 | 372 | 100.0 | 8354 | 6 ABS52100 | ABS52100 Human TEN |
| 7 | 372 | 100.0 | 8438 | 6 ABS52100 | ABS52100 Human TEN |
| 8 | 372 | 100.0 | 8645 | 6 ABS78652 | ABS78652 Human cDN |

| | | | | | | |
|----|-------|------|-------|---|----------|-----------|
| 9 | 254 | 68.3 | 8473 | 6 | ABQ82345 | Human NOV |
| 10 | 254 | 68.3 | 8487 | 6 | ABQ82346 | Human NOV |
| 11 | 254 | 68.3 | 8645 | 6 | ABQ82344 | Human NOV |
| 12 | 234 | 62.9 | 8675 | 6 | ABQ82343 | Human NOV |
| 13 | 220 | 59.1 | 12879 | 6 | ABK92230 | Prostate |
| 14 | 220 | 59.1 | 13202 | 4 | AAK51828 | Human pol |
| 15 | 195.5 | 52.6 | 1429 | 8 | ACD40264 | Human bre |
| 16 | 194.5 | 52.3 | 1430 | 5 | AAS14084 | Human FCT |
| 17 | 194.5 | 52.3 | 1430 | 5 | ADB32021 | Human FCT |
| 18 | 194.5 | 52.3 | 1431 | 4 | AAF27861 | Human NOV |
| 19 | 194.5 | 52.3 | 1727 | 7 | ADA53142 | Human cod |
| 20 | 194.5 | 52.3 | 1743 | 5 | AAS01213 | DNA encod |
| 21 | 194.5 | 52.3 | 4245 | 7 | AAI60066 | Human Pco |
| 22 | 194.5 | 52.3 | 9058 | 7 | ACC72051 | Human Pco |
| 23 | 194.5 | 52.3 | 9695 | 7 | ACC72052 | Human Pco |
| 24 | 194.5 | 52.3 | 9729 | 5 | AAS14089 | Human FCT |
| 25 | 194.5 | 52.3 | 9729 | 5 | ADB32028 | Human FCT |
| 26 | 194.5 | 52.3 | 9826 | 5 | AAS14085 | Human FCT |
| 27 | 194.5 | 52.3 | 9826 | 5 | ADB32023 | Human FCT |
| 28 | 166 | 44.6 | 1534 | 9 | ADB07179 | Novel cod |
| 29 | 149 | 40.1 | 417 | 7 | ACD05506 | CDNA enco |
| 30 | 149 | 40.1 | 643 | 5 | AAS67419 | DNA enco |
| 31 | 149 | 40.1 | 1692 | 7 | ACD05836 | Novel hum |
| 32 | 128 | 34.4 | 10242 | 4 | ABL29075 | Drosophil |
| 33 | 128 | 34.4 | 17131 | 4 | ABL29074 | Drosophil |
| 34 | 106 | 28.5 | 2569 | 4 | ABL17563 | Drosophil |
| 35 | 70.5 | 19.0 | 1467 | 3 | AAX93796 | Sequence |
| 36 | 70.5 | 19.0 | 26776 | 2 | AAZ20254 | Borrelia |
| 37 | 69 | 18.5 | 3341 | 4 | AAS27773 | DNA enco |
| 38 | 69 | 18.5 | 3341 | 4 | ADB94576 | Novel hum |
| 39 | 69 | 18.5 | 3342 | 4 | AAS27774 | DNA enco |
| 40 | 69 | 18.5 | 3342 | 4 | ADB94577 | Novel hum |
| 41 | 68 | 18.3 | 2577 | 9 | ADB63136 | Human cDN |
| 42 | 67.5 | 18.1 | 523 | 5 | AAS05481 | Mammalian |
| 43 | 67.5 | 18.1 | 748 | 2 | AAZ00430 | Human sec |
| 44 | 67.5 | 18.1 | 2105 | 4 | ABL08203 | Drosophil |
| 45 | 67.5 | 18.1 | 4105 | 4 | ABL08202 | Drosophil |

ALIGNMENTS

RESULT 1

AAV19252
ID AAV19252 standard; cDNA; 2387 BP.

XX AAV19252;

XX 17-AUG-1998 (first entry)

XX Human gamma-hergulin cDNA clone 20.

XX Gamma-hergulin; gamma-HRG; human; autocrine growth factor;
XX breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor;
XX cell proliferation; cell differentiation; cell survival;
XX neurological disorder; muscular disorder; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 307..1890

XX /*tag= a

XX FT misc_feature 1186..1263

XX /*tag= b

XX FT /*note= "insert DNA"

XX WO9802541-A1.

XX 22-JAN-1998.

XX 08-JUL-1997; 97WO-US011841.

XX 12-JUL-1996; 96US-0021640P.

PA (GETH) GENENTECH INC.
XX
PI Schaefer GM, Sliwowski M;
XX
XX WPI; 1998-110589/10.
DR P-PSDB; AAW44817.
XX
XX
PT DNA encoding gamma-hergulin - used to activate ErbB receptor and to
PT enhance proliferation, differentiation or survival of a cell.
XX
XX Disclosure; Page 56-57; 81pp; English.
XX
XX Clone 20 encodes an N-terminally truncated isoform (see AAW44818) of
XX human gamma-hergulin (gamma-HRG) (see also AAW44817), a novel member of
XX the hergulin superfamily. It was isolated from a MDA-MB-175 human breast
XX cancer cell library after screening with an EGF-like domain. Compared to
XX gamma-HRG cDNA (see AAV19251), clone 20 contains a 78 base insert between
XX codons 560 and 561, and lacks the 5' end. The truncated gamma-HRG was
XX insoluble when expressed as a thioredoxin fusion protein in Escherichia
XX coli cells. Gamma-HRG nucleic acid and polypeptide are claimed, together
XX with their uses e.g. for enhancing the proliferation, differentiation or
XX survival of glial or muscle cells
SQ Sequence 2387 BP; 641 A; 640 C; 550 G; 556 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,43e-42 Length: 2387
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AAV19252 (1-2387)

QY 1 ThrPheTrrArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 853 ACTTTCGGAGATCTCAAGTGTTCATAGACCATCTGTGCATCTGAAATTCATATGTCT 912

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 913 CTGGGAAGAGCAGCCCTGGTGGCATTTATGGCAGAAAGGCTCCCTCTTCACATACA 972

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 973 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGCTCTTAACCCAGGAGGCGCGGAGC 1032

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1033 CTAGAGGGAGCCCGCGCCAGTCTCGGGGAAC 1065

RESULT 2
AAV19251
ID AAV19251 standard; cDNA; 3111 BP.
XX
XX AAV19251;
AC
XX
DT 17-AUG-1998 (first entry)
XX
XX Human gamma-hergulin cDNA.
DE
XX Gamma-hergulin; gamma-HRG; human; autocrine growth factor;
KW breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor;
KW cell proliferation; cell differentiation; cell survival;
KW neurological disorder; muscular disorder; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 534..2645
FT /*tag= a
XX
XX WO9802541-A1.
PN

XX
PD 22-JAN-1998.
XX
PF 08-JUL-1997; 97WO-US011841.
XX
XX 12-JUL-1996; 96US-0021640P.
XX
XX (GETH) GENENTECH INC.
PA
XX Schaefer GM, Sliwowski M;
XX
XX WPI; 1998-110589/10.
DR P-PSDB; AAW44817.
XX
XX DNA encoding gamma-hergulin - used to activate ErbB receptor and to
XX enhance proliferation, differentiation or survival of a cell.
XX
XX Claim 21; Fig 1A-C; 81pp; English.
XX
XX This nucleic acid molecule codes for human gamma-hergulin (gamma-HRG)
XX (see AAW44817), a novel member of the hergulin superfamily, that has a
XX unique N-terminal domain not present in previously identified hergulins.
XX Gamma-HRG is secreted by human breast cancer MDA-MB-175 cells. It leads
XX to the formation of a constitutive active receptor complex and stimulates
XX the growth of these cells in an autocrine manner. The nucleic acid was
XX isolated by screening a MDA-MB-175 with a probe corresponding to the EGF-
XX like domain and part of the N-terminal sequences of HRG-beta-3. It can be
XX used for the recombinant production of gamma-HRG, or for in vivo or ex
XX vivo gene therapy. A claimed nucleic acid, which is complementary to the
XX nucleic acid sequence encoding the N-terminal domain of gamma-HRG, is
XX able to reduce production of gamma-HRG by MDA-MB-175 cells. The invention
XX provides claimed methods for activating an ErbB receptor and for
XX enhancing proliferation, differentiation or survival of a cell by
XX contacting the cell (preferably a glial or muscle cell) with gamma-HRG
XX
SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,03e-42 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AAV19251 (1-3111)

QY 1 ThrPheTrrArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTTCGGAGATCTCAAGTGTTCATAGACCATCTGTGCATCTGAAATTCATATGTCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1741 CTGGGAAGAGCAGCCCTGGTGGCATTTATGGCAGAAAGGCTCCCTCTTCACATACA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGCTCTTAACCCAGGAGGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGAGCCCGCGCCAGTCTCGGGGAAC 1893

RESULT 3
AAV19251
ID AAV19251 standard; cDNA; 3111 BP.
XX
XX AAV19251;
AC
XX
XX 26-OCT-1999 (first entry)
XX
XX Gamma-hergulin cDNA.
DE
XX

KW Gamma-herregulin; HRG-alpha; human; ligand; HER2; HER3; HER4; receptor;
KW lung surfactant; respiratory distress syndrome; emphysema;
KW epithelial growth factor; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 334..2640
FT /*tag= a
XX
XX WO9939729-A2.
XX
XX 12-AUG-1999.
XX
XX 03-FEB-1999; 99WO-US002390.
XX
XX 04-FEB-1998; 98US-00020598.
XX
XX (GETH) GENENTECH INC.
XX (IOWA) UNIV IOWA RES FOUND.
XX
XX Sliwkowski M, Kern JA;
XX WPI; 1999-494213/41.
XX P-FSDB; AAY06639.
XX
XX Herregulin ligands can be used to induce epithelial cell growth, and to
XX promote repair and healing of tissue damage or injury.
XX
XX Disclosure; Page 111-116; 120pp; English.
XX
XX This is the nucleotide sequence of cDNA coding for gamma-herregulin (gamma
XX -HRG, see AAY06639). The invention provides HRG ligands, including gamma-
XX HRG, that have affinity for and stimulate HER2, HER3 and/or HER4
XX receptors in autophosphorylation. A new method of treating respiratory
XX distress syndrome in humans uses HER2, HER3 and/or HER4 receptor ligands
XX as epithelial growth factors. A novel method of inducing epithelial cell
XX growth and/or proliferation comprises contacting a normal epithelial cell
XX which expresses HER2, HER3 and/or HER4 receptors with an isolated ligand
XX which activates HER2, HER3, HER4 receptors or their combination. Also
XX claimed are methods of increasing lung surfactant protein A, or of
XX treating chronic obstructive pulmonary disease, respiratory distress or
XX emphysema, by administering an effective amount of an isolated HER ligand
XX to a patient
XX
SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,03e-42 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AAX87705 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
DB 1681 ACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTCATCTGAATTCATGTGCT 1740
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
DB 1741 CTGGGAAGGCGAGCCCTGGTGGCATTTATGGCAGAAAGGGCTCCCTCTTCACATACA 1800
QY 41 GlnPheAspPheValGlnLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
DB 1801 CAGTTTGACTTTGTGGAGTGTGTGGATGGCAGAGGCTCTTAACCGAGGCGGGAGC 1860
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
DB 1861 CTAGAGGGGACCCCGCCAGTCTCGGGGA 1893

RESULT 4
RAD00791
ID AAD00791 standard; cDNA; 3111 BP.
XX
XX AAD00791;
XX
XX 21-SEP-2000 (first entry)
XX
XX Human Herregulin variant, gamma-HRG encoding cDNA.
XX
XX Herregulin; variant; gamma-HRG; human; inner-ear-supporting cell;
XX activator; HER2 receptor; HER3 receptor; HER-2 oncogene; chromosome 8p;
XX hair cell related hearing disorder; ototoxic injury; tissue damage;
XX acoustic assault; degenerative hearing loss; balance impairment;
XX treatment; hair cell; surgical injury; physical injury;
XX inner ear disorder; ss.
XX
XX Homo sapiens.
XX
XX FH Key Location/Qualifiers
FT CDS 334..2640
FT /*tag= a
FT /product= "Human Herregulin variant, gamma-HRG protein"
XX
XX WO200027426-A1.
XX
XX 18-MAY-2000.
XX
XX 28-OCT-1999; 99WO-US025744.
XX
XX 07-NOV-1998; 98US-0107522P.
XX
XX (GETH) GENENTECH INC.
XX
XX Gao W;
XX WPI; 2000-376313/32.
XX P-FSDB; AAY1176.
XX
XX Method for inducing hair cell generation and inner-ear-supporting cell
XX growth regeneration and proliferation, useful for treating hearing
XX disorders.
XX
XX Disclosure; Fig 7; 141pp; English.
XX
XX The patent discloses a method for inducing hair cell generation, or inner
XX -ear-supporting cell growth, regeneration, and/or proliferation, by
XX heregulins (HRG), the ligand for HER2/HER3 dimeric receptors. Heregulin
XX proteins function as activators of HER-2 oncogene and result from
XX alternate splicing of a single gene mapped to chromosome 8p. The two
XX major types, alpha and beta HRG's are based on two variant EGF-like
XX (epidermal growth factor) domains, that differ in their C-terminal ends.
XX This method can be used to increase the number of inner-ear-supporting
XX cells and for treatment of hair cell related hearing disorders and
XX disease states associated with tissue damage, e.g. ototoxic injury,
XX acoustic assault, degenerative hearing loss, balance impairments, damage
XX associated with surgery or physical injury and inner ear disorders
XX related to hair cell dysfunction. The present sequence is the human
XX herregulin variant, gamma-HRG protein encoding cDNA, isolated from a
XX native HRG source, such as human MDA-MB-175 cells. It has atleast one
XX biological property as the native sequence
XX
SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,03e-42 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AAD00791 (1-3111)

CC treating respiratory distress or emphysema, for treating chronic
 CC obstructive pulmonary disease (COPD) e.g. chronic bronchitis, asthma,
 CC neonatal pulmonary diseases including neonatal respiratory distress
 CC syndrome, meconium aspiration syndrome, chronic lung disease of the
 CC neonate, congenital diaphragmatic hernia and acute lung injuries
 CC including smoke or chemical inhalation, pneumonia due to aspiration,
 CC radiation, near drowning, cystic fibrosis and other epithelial cell
 CC trauma diseases including injuries associated with surgical wounds and
 CC resections, ulcers, lesions and tissue tears as normal epithelial cell
 CC growth factors. For treating infants/neonates with respiratory distress
 CC as well as youth and adult with poor lung function due to lung injury or
 CC damage. The HRG binds with varying and very high affinity to the HER2,
 CC HER3 and/or HER4 receptors. The method stimulates growth and
 CC proliferation of the epithelial cells, repairing and re-establishing the
 CC cellular barriers of organs and allowing the affected tissue to develop
 CC normal physiological functions more quickly. Hence the method improves
 CC oxygenation and speeds of the development of a barrier to infection while
 CC treating lung cells that are damaged by inhalation of smoke resulting in
 CC emphysema. It also facilitates regeneration of epithelial cells. The
 CC present sequence represents the coding sequence of human heregulin gamma-
 CC HRG
 XX
 SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,03e-42 Length: 3111
 Score: 372.00 Matches: 71
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AAS18526 (1-3111)

Qy 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
 Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCATCTGAAATCAATGTGCT 1740

Qy 21 LeuGlyLysAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
 Db 1741 CTGGAAAGGCAGCCCTGGTGGCATTTATGGCAGAAAAGCCCTCCCTTCACATACA 1800

Qy 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlnSerArgGlyThr 71
 Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGTCTTAACCCAGGAGCGCGGAGC 1860

Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
 Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGGAAC 1893

RESULT 6
 ABS52100
 ID ABS52100 standard; DNA; 8354 BP.
 AC ABS52100;
 XX
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Human TEN-M4-like gene.
 XX
 KW Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
 KW cell signal processing; metabolic pathway modulation; metabolic disorder;
 KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
 KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
 KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
 KW memory defect; infertility; congenital heart defect; hair growth;
 KW pigmentation disorder; endocrine disorder; neurological disease;
 KW gastro-intestinal disease; reproductive; respiratory disease; health;
 KW bone marrow transplantation; endocrine disease; allergy; inflammation;
 KW nephrological disorder; urinary system disorder; age-related disorder;
 KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
 KW adipocyte complement-related C1q tumour necrosis factor; out at first;
 KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;

Qy 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
 Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCATCTGAAATCAATGTGCT 1740

Qy 21 LeuGlyLysAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
 Db 1741 CTGGAAAGGCAGCCCTGGTGGCATTTATGGCAGAAAAGCCCTCCCTTCACATACA 1800

Qy 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlnSerArgGlyThr 71
 Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGTCTTAACCCAGGAGCGCGGAGC 1860

Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
 Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGGAAC 1893

RESULT 5
 AAS18526
 ID AAS18526 standard; cDNA; 3111 BP.
 AC AAS18526;
 XX
 XX
 DT 26-FEB-2002 (first entry)
 XX
 DE DNA encoding human heregulin, gamma-HRG.
 XX
 KW Human; heregulin; antiasthmatic; antiinflammatory; vulnerary; antiulcer;
 KW epithelial cell growth; HER2; HER3; HER4; HRG; lung cell; asthma;
 KW lung surfactant protein A; respiratory distress; emphysema; pneumonia;
 KW chronic obstructive pulmonary disease; COPD; chronic bronchitis; ulcer;
 KW neonatal pulmonary disease; neonatal respiratory distress syndrome;
 KW meconium aspiration syndrome; congenital diaphragmatic hernia; lesion;
 KW acute lung injury; cystic fibrosis; surgical wound; resection;
 KW growth factor; smoke inhalation; gamma-HRG; ss.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 334..2640
 FT /*tag= a
 FT /product= "Human heregulin, gamma-HRG"
 XX
 XX US2001023241-A1.
 XX
 XX 20-SEP-2001.
 XX
 XX 02-FEB-2001; 2001US-00773517.
 XX
 XX 04-FEB-1998; 98US-0073866P.
 XX
 XX 02-FEB-1999; 99US-00243198.
 XX
 XX (SLIW/) SLIWKOWSKI M X.
 XX (KERN/) KERN J A.
 XX
 XX SlwKowski MX, Kern JA;
 XX
 XX WPI; 2001-595807/67.
 XX P-PSDB; AAU09891.
 XX
 XX Inducing epithelial cell growth and/or proliferation, useful in the
 XX treatment of respiratory disease, comprises use of heregulin ligand as
 XX growth factor.
 XX
 XX Disclosure; Fig 7; 69pp; English.
 XX
 XX The invention relates to inducing epithelial cell growth and/or
 XX proliferation comprising contacting a normal epithelial cell, which
 XX expresses HER2, HER3 and/or HER4 (preferably HER2/HER3, HER2/HER4,
 XX HER3/HER4, HER3 or HER4) receptors with an isolated ligand such as
 XX heregulin (HRG) which activates HER2, HER3 and/or HER4 receptors. The
 XX method is used for inducing epithelial cell (preferably lung cell) growth
 XX and/or proliferation, for increasing lung surfactant protein A and for

type 1a membrane sushi-containing domain; butyrophilin;
type 1a membrane-sushi domain containing; SNP; gene; ds;
single nucleotide polymorphism.

Homo sapiens.

Key Location/Qualifiers
variation replace(117,G)
/*tag= a
/standard_name= "Single nucleotide polymorphism (SNP)"
replace(225,C)
/*tag= b
/standard_name= "Single nucleotide polymorphism (SNP)"
variation replace(260,A)
/*tag= c
/standard_name= "Single nucleotide polymorphism (SNP)"

WO200257453-A2.

25-JUL-2002.

19-DEC-2001; 2001WO-US050331.

19-DEC-2000; 2000US-0256704P.
20-DEC-2000; 2000US-0257314P.
02-MAY-2001; 2001US-0289153P.
29-MAY-2001; 2001US-0294075P.
24-JUL-2001; 2001US-0307506P.
10-AUG-2001; 2001US-0311590P.
10-AUG-2001; 2001US-0311613P.
29-AUG-2001; 2001US-0315617P.
14-SEP-2001; 2001US-0322358P.

(CURA-) CURAGEN CORP.

Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
WPI; 2002-590744/63.
P-PSDB; ABG70388.
Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
atherosclerosis, metabolic disorders, diabetes, obesity, infectious
disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
cancer.

Claim 8; Page 50-52; 318pp; English.

The present invention relates to new NOVX polypeptides. The invention is
useful for treating or preventing a NOVX-associated disorder such as
cardiomyopathy or atherosclerosis, where the disorder is related to cell
signal processing and metabolic pathway modulation in a subject,
preferably human. The invention is also useful for treating metabolic
disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
disorders, hematopoietic disorders and various cancers. The molecules of
the invention are also useful for treating or preventing cirrhosis,
pancreatitis, learning and memory defects, infertility, congenital heart
defects, acne, hair growth, pigmentation disorders, endocrine disorders,
respiratory disease, gastro-intestinal diseases, reproductive health,
neurological diseases, bone marrow transplantation, endocrine diseases,
allergy and inflammation, nephrological disorders, urinary system
disorders, neuropsychiatric disorders and age-related disorders. The
present nucleic acid sequence represents a NOVX gene. This sequence
encodes a NOVX protein of the invention

Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7,59e-42 Length: 8354

Score: 372.00 Matches: 71

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x ABS52100 (1-8354)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1382 ACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCATCTGAAATTCATGTGCT 1441
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyArgLysGlyLeuProSerHisThr 40
Db 1442 CTGGGAAGGAGCCCTGTTGGCAITTTATGCGAAAAGGCTCCCTCTTCATACATA 1501
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgLysLeuLeuThrGlnGluAlaArgSer 60
Db 1502 CAGTTTGTGAGTCTGCTGGATGCGAGGAGCTCTTAACCCAGAGGCGGAGC 1561
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1562 CTAGAGGGGACCCCGCCAGTCTCGGGGAAC 1594
RESULT 7
ABN85378
ID ABN85378 standard; DNA; 8438 BP.
XX
AC ABN85378;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human NOV1, TEN-M4 like protein, coding sequence.
XX
KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW TEN-M4 like protein; chromosome 11; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 4..8395
FT /*tag= a
FT /trans_except= (pos: 1138..1147,aa:Met)
FT /product= "NOV1 protein"
XX
WO200255704-A2.
XX
18-JUL-2002.
XX
09-JAN-2002; 2002WO-US000554.
XX
09-JAN-2001; 2001US-0260417P.
XX
10-JAN-2001; 2001US-0260831P.
XX
28-FEB-2001; 2001US-0272338P.
XX
09-MAR-2001; 2001US-0274876P.
XX
18-APR-2001; 2001US-0284704P.
XX
FA (CURA-) CURAGEN CORP.
XX
PI Padigar M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
PI Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;
PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
PI Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
PI Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
PI Gunther E, Stone DJ;
XX
WPI; 2002-590674/63.
P-PSDB; ABB98401.
XX
DR

XX NOVX polypeptides and encoding polynucleotides, useful for preventing or
PT treating NOVX-associated disorders e.g. cancer, inflammation, or
PT Alzheimer's disease, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Claim 9; Page 8-9; 358pp; English.
XX
XX The present sequence is a coding sequence for a NOV protein. The NOV
CC proteins and coding sequences are useful for treating or preventing NOV-
CC associated disorders or in the manufacture of a medicament for treating
CC the disorders, such as cancer, heart disease, inflammation, autoimmune
CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
CC IgA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
CC and other wasting disorders associated with chronic diseases. NOV1 is a
CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
XX
XX Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 7.7e-42 Length: 8438
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x ABN85378 (1-8438)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1358 ACTTTCTGGAGATCTCAAGTGTTTCATAGACCATCTGTGTGATCTCAAAATTCATGTGCT 1417

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 1418 CTGGGAAGCAGCCCTGGTGTGTCATTTATGGCAGAAAGCGCTCCCTCTTCACATACA 1477

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1478 CAGTTTACATTTCTGGAGCTGTGGATGGCAGGAGGCTCTAAACCCAGGAGGCGCGGAGC 1537

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1538 CTAGAGGAGGAGCCCGCCGACGTCGCGGAACT 1570

RESULT 8
ABS78652
ID ABS78652 standard; cDNA; 8645 BP.
XX
XX ABS78652;
XX
XX 16-DEC-2002 (first entry)
XX
XX Human cDNA encoding CGDD10, INCYTE 7488573CB1.
XX
XX Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
XX cell proliferative disorder; arteriosclerosis; atherosclerosis;
XX cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
XX polycythaemia vera; primary thrombocytopaenia; developmental disorder;
XX renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
XX neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
XX reproductive disorder; infertility; autoimmune disorder; gout; allergy;
XX inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
XX autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
XX diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
XX multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
XX rheumatoid arthritis.
XX
XX Homo sapiens.
XX
XX W0200272830-A2.
XX

PD 19-SEP-2002.
XX
XX 08-FEB-2002; 2002WO-US003715.
XX
XX 09-FEB-2001; 2001US-0268111P.
XX
XX 23-FEB-2001; 2001US-0271175P.
XX
XX 08-MAR-2001; 2001US-0274503P.
XX
XX 09-MAR-2001; 2001US-0274552P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
XX Ding L, Xu Y, Gletzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
XX Lu DM, Richardson TW, Tran UK, Khare R, Walia NK;
XX
XX WPI; 2002-723356/78.
XX P-PSDB; ABG97359.
XX
XX New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.
XX
XX Claim 5; Page 175-178; 181pp; English.
XX
XX The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death) a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotides and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX
XX Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;
SQ
Alignment Scores:
Pred. No.: 7.95e-42 Length: 8645
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x ABS78652 (1-8645)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1452 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCTGTGTGATCTCAAAATTCATGTGCT 1511

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 1511

KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 299..8140
XX FT /*tag= a
XX FT /product= "NOV15d"
XX
XX WO200262999-A2.
XX
XX 15-AUG-2002.
XX
XX 31-DEC-2001; 2001WO-US049976.
XX
XX 29-DEC-2000; 2000US-0258928P.
XX 02-JAN-2001; 2001US-0259415P.
XX 04-JAN-2001; 2001US-0259785P.
XX 20-FEB-2001; 2001US-0269814P.
XX 09-MAR-2001; 2001US-0279863P.
XX 29-MAR-2001; 2001US-0279832P.
XX 29-MAR-2001; 2001US-0279833P.
XX 13-APR-2001; 2001US-0283889P.
XX 18-APR-2001; 2001US-0284447P.
XX 25-APR-2001; 2001US-0286683P.
XX 29-MAY-2001; 2001US-0294080P.
XX 16-AUG-2001; 2001US-0312915P.
XX 17-AUG-2001; 2001US-0313325P.
XX 17-SEP-2001; 2001US-0322699P.
XX 26-NOV-2001; 2001US-0333350P.
XX
XX (CURA-) CURAGEN CORP.

XX Spytek KA, Li L, Wolenc AR, Vernet CM, Eisen A, Liu X;
XX PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
XX PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
XX PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
XX PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
XX WPI; 2002-732706/79.
XX P-PSDB; ABP53589.

XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
XX associated disorders, such as cancers, neurological disorders, disorders
XX of vesicular transport, gastrointestinal disorders, and autoimmune
XX diseases.

XX Claim 8; Page 123-125; 44app; English.

XX The present invention describes novel human proteins designated NOVX,
XX where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
XX cyrotatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,
XX antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
XX immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
XX protozoacide and antihelminthic activities, and can be used in gene
XX therapy. The NOVX proteins, nucleotides or antibodies can be used in the
XX manufacture of a medicament for treating a syndrome associated with a
XX human disease selected from NOVX-associated disorder, such as cancers
XX (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
XX ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
XX ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
XX disease), disorders of vesicular transport (e.g. cystic fibrosis,
XX diabetes mellitus, Grave's disease, or goitre), gastrointestinal
XX disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
XX autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
XX anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
XX and protozoal infections. The NOVX proteins can be used as immunogens to
XX produce antibodies and as vaccines. The NOVX nucleotide sequences may be

CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15d, which is
CC located on chromosome 4

XX Sequence 8487 BP; 2306 A; 2029 C; 2148 G; 2001 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 5, 38e-25 Length: 8487
Score: 254.00 Matches: 48
Percent Similarity: 80.88% Conservative: 7
Best Local Similarity: 70.59% Mismatches: 13
Query Match: 68.28% Indels: 0
DB: Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x ABQ82346 (1-8487)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
DB 1226 TTCTGGAGATCACAGCTCTTCATGATCAGCCACAGATTTCTTTAAATTCATATCTCTT 1285
QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln 41
DB 1286 CAGAAGGATGCATTGATTGGAGTATAGCCGAAAGGTTACCGCTTCCCATATCTAG 1345
QY 42 PheAspPheValIleLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61
DB 1346 TATGACTTCGTGGAGCTCTGGATGCGCAGCAGCTGATTGCCAGAGAGCGAGCGCTG 1405
QY 62 GluGlyThrProArgGlnSerArg 69
DB 1406 CTTGAGACGAGAGAGCGCGCGG 1429

RESULT 11
ABQ82344
ID ABQ82344 standard; cDNA; 8645 BP.

XX ABQ82344;

XX 17-DEC-2002 (first entry)

XX Human NOV15b encoding cDNA SEQ ID NO:37.

XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
XX cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;
XX antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
XX antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
XX cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
XX stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
XX Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
XX diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
XX ulcerative colitis; gastric disorder; duodenal disorder; infection;
XX autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
XX rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 151..8316
XX FT /*tag= a
XX FT /product= "NOV15b"

XX WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

XX 29-DEC-2000; 2000US-0258928P.

XX 02-JAN-2001; 2001US-0259415P.

XX 04-JAN-2001; 2001US-0259785P.

XX 20-FEB-2001; 2001US-0269814P.

XX 09-MAR-2001; 2001US-0279863P.

Db 1493 GATGATACACAGCACTCCCTCGG 1516

RESULT 14

AAK51828

ID AAK51828 standard; cDNA; 13202 BP.

XX

AC AAK51828;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 373.

XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation; ss.

XX

OS Homo sapiens.

XX

FN WO200157190-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US004098.

XX

PR 03-FEB-2000; 2000US-00496914.

PR 27-APR-2000; 2000US-00560875.

PR 20-JUN-2000; 2000US-00598075.

PR 19-JUL-2000; 2000US-00620325.

PR 01-SEP-2000; 2000US-00654936.

PR 15-SEP-2000; 2000US-00663561.

PR 20-OCT-2000; 2000US-00693325.

PR 30-NOV-2000; 2000US-00728422.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX

DR WPI; 2001-476283/51.

DR P-PSDB; AAM78695.

XX

PT Nucleic acids encoding polypeptides with cytokine-like activities, useful

PT in diagnosis and gene therapy.

XX

PS Claim 1; Page 1414-1426; 6221pp; English.

XX

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

CC sequence listing were missing at the time of publication

XX

SQ Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.91e-20 Length: 13202

Score: 220.00 Matches: 36

Percent Similarity: 79.41% Conservative: 18

Best Local Similarity: 52.94% Mismatches: 14

Query Match: 59.14% Indels: 0

DB: 4 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AAK51828 (1-13202)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21

Db 1639 TTCTGGCGTTTCCAGATTACTATCCACCATCAATATATATCTGAAGTTCAATATTTCTTTA 1698

QY 22 GlyLysAlaAlaLeuValGlyIleTrpGlyArgLysGlyLeuProProSerHisThrGln 41

Db 1699 GCCAAGGACTCTCTGCTGGGAATTTATGCGACAGAGAAACATTCCACCTACATCTCAG 1758

QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61

Db 1759 TTTGATTTTGTAAACTAATGATGGCAACAGCTGGTCAACAGGACTCCCAAGGGCTCT 1818

QY 62 GluGlyThrProArgGlnSerArg 69

Db 1819 GATGATACACAGCACTCCCTCGG 1842

RESULT 15

ACD40264

ID ACD40264 standard; DNA; 1429 BP.

XX

AC ACD40264;

XX

DT 03-SEP-2003 (first entry)

XX

DE Human breast tumour associated protein 47-like polypeptide NOV11 DNA.

XX

KW Tissue typing; cancer; breast cancer; colon cancer; lung cancer; sarcoma;

KW pancreatic cancer; uterine cancer; organ transplantation disorder; ds;

KW cardiovascular disease; melanoma; atherosclerosis; diabetes mellitus;

KW ischaemic heart disease; haemorrhage; peripheral vascular disease; gene;

KW thrombosis; hypertension; systemic lupus erythematosus; haematopoiesis;

KW tissue regeneration; wound healing; hyperproliferative disorder;

KW psoriasis; neural disorder; Parkinson's disease; Alzheimer's disease;

KW Huntington's disease; multiple sclerosis; amyotrophic lateral sclerosis;

KW ALS; peripheral neuropathy; nervous system tumour; neurotoxin; tremor;

KW neuropathy; acute brain injury; peripheral nerve trauma; human; NOVX;

KW gene therapy; epilepsy; breast tumour associated protein 47.

XX

OS Homo sapiens.

XX

FN US2003027158-A1.

XX

PD 06-FEB-2003.

XX

PF 15-OCT-2001; 2001US-00977418.

XX

PR 03-JUN-1999; 99US-0137322P.

PR 16-MAR-2000; 2000US-0189810P.

PR 22-MAR-2000; 2000US-0191158P.

PR 30-MAR-2000; 2000US-0193086P.

PR 03-MAY-2000; 2000US-0201388P.

PR 31-MAY-2000; 2000US-00584411.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Shimkets RA, Fernandes E, Herrman J, Vernet C;

XX

DR WPI; 2003-492028/46.

DR P-PSDB; ABO23243.

XX

PT New nucleic acid sequence encoding a human breast tumor-associated

PT protein 47-like polypeptide, useful for treating cardiovascular

PT disorders, neural disorders, diabetes mellitus and cancers.

XX

PS Disclosure; Page 29-30; 100pp; English.

XX

CC The invention relates to a new isolated NOV4 nucleic acid. The nucleic

CC acid is useful for identifying a compound that binds the nucleic acid.

CC The nucleic acid is useful in gene therapy, in screening assays, in

CC detection assays e.g. chromosomal mapping, cell and tissue typing and

CC forensic biology, predictive medicine e.g. diagnostic assays, prognostic

CC assays, monitoring clinical trials and pharmacogenomics and methods of

CC treatment including therapeutic and prophylactic. The nucleic acid is
CC also useful for expressing NOVX protein. The nucleic acid is also useful
CC to provide polynucleotide reagents e.g. labelled probes that are useful
CC in an in situ hybridisation technique, for identifying a specific tissue
CC (for example brain tissue) and for use in forensic science. The nucleic
CC acid is also useful for mapping genes on a chromosome and thus locating
CC gene regions associated with genetic disease, identifying an individual
CC from a minute biological sample and to aid in forensic identification of
CC biological sample. The nucleic acid is also useful for treating cancer,
CC especially cancers of the breast, colon, lung, pancreas or uterus, or a
CC melanoma or sarcoma. The nucleic acid is also useful for treating
CC disorders related to organ transplantation, cardiovascular diseases,
CC atherosclerosis, ischaemic heart disease, haemorrhage, diabetes mellitus,
CC peripheral vascular disease, thrombosis, hypertension and systemic lupus
CC erythematosus. NOVX protein encoded by the nucleic acid is useful for
CC regulating haematopoiesis, for regeneration of bone, cartilage, tendon
CC ligament and/or nerve tissue growth or regeneration and for wound
CC healing. The nucleic acid is also useful for treating infections,
CC hyperproliferative disorders e.g. psoriasis, and neural disorders
CC including Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC multiple sclerosis, amyotrophic lateral sclerosis (ALS), peripheral
CC neuropathy, tumours of the nervous system, exposure to neurotoxins, acute
CC brain injury, peripheral nerve trauma or injury and other neuropathies,
CC epilepsy, and/or tremors. The present sequence represents DNA encoding a
CC human breast tumour associated protein 47-like polypeptide
XX
SQ Sequence 1429 BP; 349 A; 426 C; 340 G; 312 T; 0 U; 2 Other;

Alignment Scores: 1-12e-17 Length: 1429
Pred. No.: 195.50 Matches: 38
Score: 70.59% Conservative: 10
Percent Similarity: 55.88% Mismatches: 15
Best Local Similarity: 52.55% Indels: 5
Query Match: 8 Gaps: 1
DB:

US-10-029-020-14_COPY_450_520 (1-71) x ACD40264 (1-1429)

| | | | |
|----|------|--|------|
| QY | 2 | PheTyrArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu | 21 |
| Db | 884 | TTTTGGAGGTCAAAATTCACATCAGTCAGCCCCAGTTCTTAAAGTTCAACATCTCCCTC | 943 |
| QY | 22 | GlyValAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThrGln | 41 |
| Db | 944 | GGGAGGAGCCCTCTCTTTGGTGTTCATTAAGAGAGGACTTCACCATCTCATGCCAG | 1003 |
| QY | 42 | PheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSerLeu | 61 |
| Db | 1004 | TATGACTTCATGGAACCTCTGGACGGGAG-----GAGNAGTGGAGTGTG | 1048 |
| QY | 62 | GluGlyThrProArgGlnSerArg | 69 |
| Db | 1049 | GTTGAGTCTCCCGGAACGCCGG | 1072 |

Search completed: August 14, 2004, 02:28:55
Job time : 135.435 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 1786.37 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_750_850
Perfect score: 617
Sequence: 1 WGAACDQACHPRCAHGT.....DISMETACGDSKNDGDGLV 101

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -Qfmt=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn 1 1 19065 @runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|-----------|-------------|
| 1 | 617 | 100.0 | 8354 | 6 | AX556500 | Sequence |
| 2 | 617 | 100.0 | 8645 | 6 | AX600210 | Sequence |
| 3 | 605 | 98.1 | 8585 | 10 | AB025413 | Mus muscu |
| 4 | 579 | 93.8 | 8438 | 6 | AX675551 | Sequence |
| 5 | 541 | 87.7 | 8964 | 10 | AB025412 | Mus muscu |
| 6 | 540.5 | 87.6 | 9722 | 10 | AF059485 | Mus muscu |
| 7 | 515 | 83.5 | 8473 | 6 | AX662357 | Sequence |
| 8 | 515 | 83.5 | 8645 | 6 | AX662355 | Sequence |
| 9 | 515 | 83.5 | 8675 | 6 | AX662353 | Sequence |
| 10 | 503.5 | 81.6 | 9264 | 5 | AB026980 | Danio rer |
| 11 | 483.5 | 78.4 | 3394 | 9 | AK125869 | Homo sapi |
| 12 | 483.5 | 78.4 | 7815 | 10 | AF135418 | Mus muscu |
| 13 | 478.5 | 77.6 | 8487 | 6 | AX662359 | Sequence |
| 14 | 467.5 | 75.8 | 8373 | 10 | AB025410 | Mus muscu |
| 15 | 464.5 | 75.3 | 8297 | 9 | AF100772 | Mus muscu |
| 16 | 454.5 | 73.7 | 8816 | 5 | AB026979 | Danio rer |
| 17 | 454 | 73.6 | 816 | 10 | AF086610 | Rattus no |
| 18 | 454 | 73.6 | 1011 | 10 | AF086609 | Rattus no |
| 19 | 454 | 73.6 | 7781 | 9 | AB032953 | Homo sapi |
| 20 | 454 | 73.6 | 9826 | 6 | AX250008 | Sequence |
| 21 | 452 | 73.3 | 8409 | 5 | GGA279031 | Gallus ga |
| 22 | 448.5 | 72.7 | 7713 | 9 | HSB08325 | Sequence |
| 23 | 448.5 | 72.7 | 7713 | 9 | HSB08325 | Sequence |
| 24 | 448.5 | 72.7 | 8689 | 6 | AX250067 | Sequence |
| 25 | 448.5 | 72.7 | 8689 | 10 | AF086607 | Rattus no |
| 26 | 448.5 | 72.7 | 8797 | 6 | AX250063 | Sequence |
| 27 | 448.5 | 72.7 | 8797 | 6 | AX250066 | Sequence |
| 28 | 448.5 | 72.7 | 8797 | 10 | AB025411 | Mus muscu |
| 29 | 448.5 | 72.7 | 9729 | 6 | AX250013 | Sequence |
| 30 | 447.5 | 72.5 | 8575 | 6 | AX21803 | Sequence |
| 31 | 435.5 | 70.6 | 8118 | 5 | GGA238613 | Gallus ga |
| 32 | 330.5 | 53.6 | 798 | 10 | AF086608 | Rattus no |
| 33 | 301 | 48.8 | 112872 | 2 | AP000716 | Homo sapi |
| 34 | 301 | 48.8 | 165812 | 2 | AP001141 | Homo sapi |
| 35 | 301 | 48.8 | 168308 | 2 | AC084775 | Homo sapi |
| 36 | 301 | 48.8 | 170156 | 2 | AP002412 | Homo sapi |
| 37 | 301 | 48.8 | 175464 | 2 | AP002887 | Homo sapi |
| 38 | 301 | 48.8 | 186084 | 9 | AP002768 | Homo sapi |
| 39 | 300 | 48.6 | 245722 | 2 | AC120288 | Rattus no |
| 40 | 300 | 48.6 | 257693 | 2 | AC103323 | Rattus no |
| 41 | 296 | 48.0 | 10471 | 3 | AF008228 | Drosophila |
| 42 | 296 | 48.0 | 11451 | 3 | DMTENN | Drosophila |
| 43 | 290 | 47.0 | 226256 | 10 | AC073599 | Mus muscu |
| 44 | 279.5 | 45.3 | 10216 | 3 | DMTENA | Drosophila |
| 45 | 278 | 45.1 | 14045 | 3 | DMODZ4 | Drosophila |

ALIGNMENTS

RESULT 1

AX556500
LOCUS AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Gangolli, E.A., Patturajan, M., Vernet, C.A., Malyankar, U.M.,
Kekuda, R., Stone, D.J., Anderson, D., Shimkets, R.A., Burgess, C.E.,
Zerhusen, B.D., Liu, X., Spytek, K.A., Casman, S.J., Boldog, F.L.,
Smithson, G., Li, L., and Ji, W.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
FEATURES
source
1. .8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6.06e-43 Length: 8354
Score: 617.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_750_850 (1-101) x AX556500 (1-8354)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
DB 2282 TGGATGGGGGACGCTCGACACCGGGCTGCGCTGGTGAATGGCAATGGCAACGCGAGA 2341
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTTPAsnGlyGluHisCysThrIleAla 40
DB 2342 TGGCGGACGGCAAGTCGAGTGCAGCCCTGGCTGGTGAATGGCAATGGCAACGCGAGA 2401
QY 41 HisTyrLeuAspArgValValLysGluCysProGlyTTPAsnGlyGluHisCysThrIleAla 60
DB 2402 CACTATCTGATAGGGTAGTTAAAGAGGGTGGCTGGTGAATGGCAATGGCAACGCGAGA 2461
QY 61 CysThrLeuAspLeuAsnGlyTTPHisCysValCysGlnLeuGlyTTPArgGlyAlaGly 80
DB 2462 TGTACCTTAGACCTGAATGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGC 2521
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
DB 2522 TGTGACACTTCATGGAGACTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGC 2581
QY 101 Val 101
DB 2582 GTG 2584
RESULT 2
LOCUS AX600210 8645 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Yue, H., Yao, M.G., Ison, C.H., Lu, Y., Warren, B.A., Elliott, V.S.,
Baughn, M.R., Ding, L., Xu, Y., Gietzen, K.J., Tang, T.Y., Lal, P.G.,
Duggan, B.M., Burford, N., Lu, D.A., Richardson, T.W., Tran, U.K.,
Khare, R. and Wallia, N.K.
TITLE Proteins associated with cell growth, differentiation, and death
JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1. .8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="inocyte ID No: 7488573CB1"
ORIGIN
Alignment Scores:
Pred. No.: 6.26e-43 Length: 8645
Score: 617.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_750_850 (1-101) x AX600210 (1-8645)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
DB 2352 TGGATGGGGGACGCTCGACACCGGGCTGCGCTGGTGAATGGCAATGGCAACGCGAGA 2411
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTTPAsnGlyGluHisCysThrIleAla 40
DB 2412 TGGCGGACGGCAAGTCGAGTGCAGCCCTGGCTGGTGAATGGCAATGGCAACGCGAGA 2471
QY 41 HisTyrLeuAspArgValValLysGluCysProGlyTTPAsnGlyGluHisCysThrIleAla 60
DB 2472 CACTATCTGATAGGGTAGTTAAAGAGGGTGGCTGGTGGTGGTGGTGGTGGTGGTGGC 2531
QY 61 CysThrLeuAspLeuAsnGlyTTPHisCysValCysGlnLeuGlyTTPArgGlyAlaGly 80
DB 2532 TGTACCTTAGACCTGAATGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGC 2591
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
DB 2592 TGTGACACTTCATGGAGACTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGC 2651
QY 101 Val 101
DB 2652 GTG 2654
RESULT 3
LOCUS AB025413 8585 bp mRNA linear ROD 08-MAY-1999
DEFINITION Mus musculus mRNA for Ten-m4, complete cds.
ACCESSION AB025413
VERSION AB025413.1 GI:4760781
KEYWORDS Ten-m4.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Ohashi, T., Zhou, X., Peng, K., Richter, B., Moergelin, M., Perez, M.T.,
Su, W., Chiquet-Ehrismann, R., Rauch, U. and Faessler, R.
TITLE Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
JOURNAL proteins expressed in many tissues
AUTHORS J. Cell Biol. (1999); In press
TITLE 2 (bases 1 to 8585)
JOURNAL Ohashi, T.
AUTHORS Direct Submission
TITLE Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University
JOURNAL Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
FEATURES (E-mail: ohashi@cc.okayama-u.ac.jp, Tel:+81-86-235-7128,
Fax:+81-86-232-7768)
Location/Qualifiers

US-10-029-020-14_COPY_750_850 (1-101) x AB025413 (1-8585)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCyHisProArgCysAlaGluHisGlyThr 20
Db 2441 TGGATGGGGCCCGCATCGACCAACGGGCTGCACCACGTCGTGTGAGAACACGGGACC 2500

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2501 TGGCGGGACGGCAAGTGGGAATCACGCCCGGTGGAATGGAGAGCAGTGACCATCGCT 2560

QY 41 HisTyrlieuAspArgValValllysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2561 CACTATCTGATAGGCTAGTTAAAGAGGGCTGCTCTGCTGTGCAATGGAATGGACG 2620

QY 61 CysThrIleuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2621 TGTAACCTGACCTGAACTGGTGGCACTGTGTCTGCAGCTGGGCTGGGAGGACCTGGC 2680

QY 81 CysAspThrSerMetGluThralaCysGlyAspSerLysAsnAspGlyAspGlyLeu 100
Db 2681 TGGCAGACATCCATGGAAACGGGCTGTGGAGACGGCAAGCAACACGAGATGCTTG 2740

QY 101 Val 101
Db 2741 GTG 2743

RESULT 4
LOCUS AX675551 8438 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 1 from Patent WO02055704.
ACCESSION AX675551
VERSION AX675551.1 GI:293333552
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Padigar, M., Li, L., Zerhusen, B.D., Casman, S.J., Shenoy, S., Spytke, K.A., Zhong, M., Gangoli, E.A., Burgess, C.E., Patturajan, M., Vernet, C.A., Taylor, S., Thornev, V.T., Miller, C.E., Guo, X., Bollog, F.L., Grosse, W.M., Alsobrook, J.P., Gerlach, V., Edingermark, S., Rotherberg, M.E., Ellerman, K., Macdougall, J., Malynkar, U., Millet, I., Peyman, J., Smithson, G., Gunther, E. and Stone, D.J.

TITLE Proteins, polynucleotides encoding them and methods of using the same
JOURNAL Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1..8438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.06e-39 Length: 8438
Score: 579.00 Matches: 98
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 93.33% Mismatches: 2
Query Match: 93.84% Indels: 4
DB: 6 Gaps: 2

US-10-029-020-14_COPY_750_850 (1-101) x AX675551 (1-8438)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCyHisProArgCysAlaGluHisGlyThr 20
Db 2249 TGGACGGGCCAGCTGTAAATCAGAGACCTGCCACCCCCTGTGGCAGACACGGGACC 2308

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 39

Db 2309 TGCCGCGACGCAAGTGGAGTGCAGCCCTGGCTGAATGGCAACACTGCACCATCTCC 2368

Qy 40 ---AlaHisTyrLeuAspArgValVallys-----GluGlyCysProGlyLeuCysAsn 56
|||||
Db 2369 CTAGCTCACTATCTGATAGGCTAGTAACTTTTCAGAGGGTGGCCCTGGTGTGCAAT 2428
|||||

Qy 57 GlyAsnGlyArgCysThrLeuAspLeuAnGlyTyrHisCysValCysGlnLeuGlyTyr 76
|||||

Db 2429 GGCACGCGCAGATGACCTTACAGCTGAATGGTGGCCTGGCTGCGCAGCTGGGCTGG 2488
|||||

Qy 77 ArgGlyValGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAsp 96
|||||

Db 2489 AGAGAGAGTGGCTGTGACACTTCCATGGAGACTGCTGCGGTGACAGCAAGACATGAT 2548
|||||

Qy 97 GlyAspGlyLeuVal 101
|||||

Db 2549 GGAGATGGCGCTGGTG 2563
|||||

RESULT 5
AB025412
LOCUS
DEFINITION Mus musculus mRNA for Ten-m3, complete cds.
ACCESSION AB025412
VERSION AB025412.1 GI:4760779
KEYWORDS Ten-m3.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Ohashi, T., Zhou, X., Feng, K., Richter, B., Moergelin, M., Perez, M.T.,
Su, W., Chiquet-Ehrismann, R., Rauch, U. and Faessler, R.
TITLE Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
J. Cell Biol. (1999) In press
REFERENCE 2 (bases 1 to 8964)
Ohashi, T.
Direct Submission
Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
(E-mail: ohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128,
Fax: +81-86-222-7768)
FEATURES
Location/Qualifiers
1..8964
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1..8964
/gene="ten-m3"
50..8197
/gene="ten-m3"
/codon_start=1
/product="Ten-m3"
/protein_id="BA077398.1"
/db_xref="GI:4760780"
/translation="MDVKRRPYCSLTSSRRKERRRYTNSSADNEECRVPTQKSYSSS
ETLKAPDHYSRLLYGNRVKDIYVREADEYTRQGNFTLRQLGVCESATRRGVAFCAE
MGLPHRGYSISAGSDADTENEAVMSPEHAMRLMGKGVKGRSSCLSSNSALTLTDT
EHNRSDESSEOPSNNGOPTLOPLPESHKQHPAHHPSITSLNRSLNTRNQSAP
PAALPAELQTTPEVSQVQLQSWLVCSNVPLESRHFLPKTGTGTPPLESTATPGVTMAG
SVSPPTREPLNTLSRSAPFKSKSVCSWRCTALCAVGSVLLAILLSYFIAMHLF
GLNHLQQTENDTFENGNSKVTPTNTVSLPSGDNKLGFGFTHEINTDSDGLDGR
GRLLIAREORNLVESRAGROARSVSLHEAGFTQYLDSGIWHLAFYNDGKNPQSFN
TIVLESVCPNCHNGCVSCTCHCFGLGDCSRAACPVLCSGNGOYSGKRCCLC
PSGKWTGECVPTQCIDPQCGRGICIMGSCACNSGYKGENCEADCLDPCGNSHG
CHGECNCPGNSCEILKTMCAQCSCGHYLRSGSCCTCDPNTWTPDSCSNEICS
VDCSGVCMGSCRCCEBGTWGPACNRACHPRCAEHGTCKDGKCSQSGWNEHCTI
AHYLDKIVKEGCPGLCNSGRCTLDQNGWHVCVQPGNRRGAGCDVAMETLCTDSKDN

ORIGIN

Alignment Scores:
Pred. No.: 1,956-36 Length: 8964
Score: 541.00 Matches: 85
Percent Similarity: 91.09% Conservative: 7
Best Local Similarity: 84.16% Mismatches: 9
Query Match: 87.68% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x AB025412 (1-8964)

Qy 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||
Db 2153 TGGACCGCGCCGCGTGTAAATCAGAGAGCTTGCACCCCTCGCTGTCTGACGACGAGC 2212
|||

Qy 21 CysArgAspGlyLysCysGluCysSerProGlyTyrAsnGlyGluHisCysThrIleAla 40
|||
Db 2213 TGCACGACGCGCAAGTGCAGTGCACCAAGATGACGAGAGCAGCTGCACAAATTGCT 2272
|||

Qy 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
|||
Db 2273 CACTATTGGATAAGATAGTTAAAGAGGGTTGACCGCGCTTGTGCAACAGCAATGGGAG 2332
|||

Qy 61 CysThrLeuAspLeuAnGlyTyrHisCysValCysGlnLeuGlyTyrArgGlyValGly 80
|||
Db 2333 TGCACACTGCACAAACCGCTGGCACTGCTGTTGCCAGCGGTGGAGAGAGCAGGC 2392
|||

Qy 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
|||
Db 2393 TGTGACGTGACCAACCGCTGGCACTGCTGTTGCCAGCGGTGGAGAGAGCAGGC 2452
|||

Qy 101 val 101
|||
Db 2453 ATT 2455
|||

RESULT 6
AF059485
LOCUS AF059485 9722 bp mRNA linear ROD 15-AUG-1998

```
DEFINITION Mus musculus Doc4 (Doc4) mRNA, complete cds.
ACCESSION AF059485
VERSION AF059485.1 GI:3170614
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS Wang, X.-Z., Kuroda, M., Sok, J., Batchvarova, N., Kimmel, R., Chung, P.,
Zinszner, H. and Ron, D.
TITLE Identification of novel stress-induced genes downstream of chop
JOURNAL EMBO J. 17 (13), 3619-3630 (1998)
MEDLINE 98315054
PUBMED 9649432
REFERENCE
AUTHORS Wang, X.-Z. and Ron, D.
TITLE Direct Submission
JOURNAL Submitted (14-APR-1998) Skirball Institute, New York University
Medical Center, 550 First Ave., New York, NY 10016, USA
FEATURES
source
1..9722
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH/3T3"
/db_xref="taxon:10090"
/cell_line="NIH-3T3"
/note="tunicamycin-treated"
1..9722
/gene="Doc4"
/note="downstream of CHOP 4"
583..9060
/gene="Doc4"
/note="similar to Drosophila melanogaster tenm/odz and
human gamma-heretulin; type II transmembrane protein"
/codon_start=1
/product="DOC4"
/protein_id="AAC31807.1"
/db_xref="GI:3170615"
/translation="MDVKKPKYSLRRDAERYTSSADSEEGKPGKQSSST
LKAYDQRLAYSRVMDVMPQEAEEFCRTGNTFLRELGLGEMTPHGLTYLTDIGL
PHGYSNGASSDADLEADTVLSPHPVLRWGRSTRSGRSCLSRRANSNLTDITDEH
NTEGAPLHCSASSTPIEQSPSPPPANESORRLGNGVQAPTPDSSEFEFVN
SFLVKSGLSLVAANDHPSLQNHPLRTPPLPHAHTPNQHAASINSLRNGT
PRNSPAPTDHSLSGEPAGSAGEPHADQNVNLNSKIPIETRNKGQPLGTWQNT
LIEMDIPASRRGAYSDGHFFKPGTSELCFTTSPGYELTSTVSPPLPRST
FSRPAFLKPSKYCNMKCAALSALISATLIVLLAYFVAMHLFNLHQLPMQEQM
MYETITDASSVPVTDVSLYPSGGTLETDRKKGAAEKPSSLFPEDSFIDSGEI
DVGRASQKTPPGTFWRSQVFIQHPVHLKFNLSLGAALVGIYGRKGLPPSHQTFD
ELDGRRLLTQEARSLGEPQORQSPVPSHSETGFTQYLDLSGIIWHLAFYNDKSEY
VSFLTIAESVDNCPNCGNDCISCTCHCFGLGFLGPDGRASCPLVCSNGQYMKG
RCLCHSGWKAECDPVINOQIDVACSHGTCIMGTCLCNFGYKGESEEDVMDPTCS
SRGVCRGECVSGVGTGNCETPRATCDQCSGHGTFLEPTGLCNQDPSWTGDCSI
EICAADCGHGVCGVGTCECDMGMAQDQACRACHRGTCRDGKCECTPGWGE
HCTIEGCLNCGNCRLLDNGWHCVCLGWRGTGDTSMETGCGDKNDGDGLVD
CMPDPCLOPLCHNPLCLGSPDLDIIQETQAPVSOQLNSFYDRILKELVGRDSTHS
IPGENPDGGAACVIGSQWMTSDGTPLGVNIFINPLPGYIISQDGSFDLVINGG
ISILPERAPITQHTLMLPWDRFFVMEIIVRHEENIEIPRDLNSFARPNPVSP
SPUTSPASSCAEGPIPEIQALOEIIVTAGCKMRLSYLSSPTPGYKSVVRIISLTP
IPLNMLHLMWAVEGLRFRKFAAAPDLSSYFIWDKTDVYNQKFLGEAFVSGVE
YESCPDLILWEKATVLOQVEIDAKSLGSLDKHNLNIOGILHKGNGENOFVSQO
PVPVIGSMNGRRRSISCPNCLGADGNKLLAPVALTCGSDGSLYVQDNYIPIEPS
GNVTNILEMSHPAHKYLATDPMGAVFLSDTNSRPVFKVSTTVKRLVKNSEVVA
GTGDCQLPEDTCGGGKATEATLNPGRITVDKGLIFYDGTWIRLVQNGIIST
LIGSNLTSARPULSCDSMEISQVRLWEPDILALNPMNSLYLDNNVLIQISENHQV
RIVAGRPMHCQVPGIDQFLSLKVAIHATLESALAVSHNGVLIYIAETDEKKNIRQ
VITTSIRIVAGAPSCDCNDANCDFSGDGYAKOAKLNTSPSLAVACDGLYVAD
LGNIRFRIRKKNKPFNTQMYELSSPIDQLYLPDTSKHLVTQSLPGDLYNRYT
TDGDTLTHIDNNNMVNRDSTGMPLMLVDPGVVYVWMTGNALSASVTQGHLE
AMWTHYNGSLGATKSNENGWTFYEDFSFRLTNVTFPGQVSSFRSDTSSVHVQ
ETSSKDDVTITHLSGGAPFILLQDVRNSYIIGADSLRLLANGMEVALQTEPH
LAGTVNPTQVGTPIIDNGLNLEWRQEQARQVTFGPRLRVNRNLISLDFD
RVTRTEKIYDHRKFTLRILYDQGRPSFWSPPSSRLNGVNVTVSPGHIAGIQRGMS
```

```
ERMEDYDQARITSRIFADCKMWSYTYLEKSMVYLHLSQRQYIFEFDKNDRLSSVTWPN
VARQLETIRSVGYRNIYQPPEGNASVIQDTECHLLHTFYLCTGRRVIVYKGLS
KLAETLYDTTKVSFTYDETAKMLKTIVNLQNEGTCTIRYROIQPLIDRQIFPTEGM
VNAREYVNDNSFRVTSQAVINETPLPIDLYRYDDVSGKTEQFKFGVIYDINQII
TTAVNTHSKFDAYGRMEQVETFRSLMYMTVOYDNNRGRVKKELKVPYANTRY
SYEDDAQLOTVSINDKPLWRSYDNLNGLHLLSPGNSARLTPLRYDLDRITLIGD
VOYKDEDSLRQGGDVFEYNSAGLIIKAYNRASGVVRVYDGLGRRVSSKSHSH
HLQFFYADLTNPTKVTLYSHSSSEITSLYDLOGLHFAELSSGDEFVACDNIGTP
LAVFSGTGLMKIQLIYAYGEIYMDTNFQIIIGYHGLYDPLTKLVHMGPRDYDL
AGRVSPDHELWKLSSNIVPEHLYMEKNPNISNSQDIKCFMDVNSWLLTFGFL
HNVLPGYKPDTDAMEPSYELVHQMKTQEWDSKSLIGVCEVQKQLKAVFTLERFD
QLYGSTISCCQAPETKKFASGSIIFGKVPALKDGRVTTDIISVANDGERRIAIL
NNAHLENLHTIDGVTHTFVKPSEGLAILLGSRRRTLENGVNVTVSQINML
SGTRRTYDIIQOYRALCLNTRYGTTVDEKRVLELARQVRAQVAREOQRLRGE
EGLRAWTDGKQVLTGRVQYDGFVTSVEQYPELSDSANNIHFMEQSEMR"
1813..1881
/gene="Doc4"
/note="putative; transmembrane-region site"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,32e-36 Length: 9722
Score: 540.50 Matches: 88
Percent Similarity: 88.12% Conservative: 1
Best local Similarity: 87.13% Mismatches: 3
Query Match: 87.60% Indels: 9
DB: 10 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x AF059485 (1-9722)

Qy 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 3046 TGGATGGGGCGCGCATGCACCAACGGCGCTGCCACCGCTGTGAGAACACGGGACC 3105
Qy 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 3106 TGGCGGAGCGGCAAGTGGAGATGACCCCGCGCTGGAGATGGAGAGCACTGCACCATC 3162
Qy 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 3163 -----GAGGGCTGCTGCTGGCTGTGCATGGAAATGGCAGA 3198
Qy 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 3199 TGTACCTTGGACCTGAATGGGTGGCACTGTCTGTCCAGCTGGCTGGCGAGGCACTGGC 3258
Qy 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 3259 TGGACACATCCATGGAAACGGCTGTGGAGACGGCAAGGACACGACGAGATGGCTTG 3318
Qy 101 Val 101
Db 3319 GTG 3321
```

RESULT 7

```
AX662357
LOCUS AX662357
DEFINITION Sequence 39 from Patent WO02062999.
ACCESSION AX662357
VERSION AX662357.1 GI:29163218
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
Malyankar, U., Shankets, R.A., Tchernev, V.T., Spaderina, S.K.,
Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
Millet, I. and Macdougall, J.R.
PROTEINS AND NUCLEIC ACIDS ENCODING SAME
```

JOURNAL Patent: WO 02062999-A 39 15-AUG-2002;
Curagen Corporation (US)

FEATURES
source
1. .8473
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 3.04e-34 Length: 8473
Score: 515.00 Matches: 84
Percent Similarity: 85.05% Conservative: 7
Best Local Similarity: 78.50% Mismatches: 10
Query Match: 83.47% Indels: 6
DB: Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x AX662357 (1-8473)

| | | | |
|----|------|---|------|
| Qy | 1 | TrpMetGlyAlaAlaCysAspGlnAtqAlaCysHisProArgCysAlaGluHisGlyThr | 20 |
| Db | 2082 | TGACGGGCGCCAGCCTGTATCAAGAGAGCTGCACCCCGCTGTGCCGAGCAGGGACC | 2141 |
| Qy | 21 | CysArgAspGlyLyScysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla | 40 |
| Db | 2142 | TGCAAGATGGCAAGTGTGAATGCAGCCAGGCTGGAATGCAGAGCACTGCCTATCGCT | 2201 |
| Qy | 41 | HisTyrLeuAspArgValVal-----lysGluGlyCysProGlyLeu | 54 |
| Db | 2202 | CACATATTTGGATAAGATAGTGTAAAGACACAGATAGGATATAAAGAGGGTGTCTCTGGTGTG | 2261 |
| Qy | 55 | CysAsnGlyAanGlyAiqCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu | 74 |
| Db | 2262 | TGCAACAGCAATGGAGAGATGTACCTGGACCACAAATGGCGAGCATTTGTGTGCCAGCCT | 2321 |
| Qy | 75 | GlyTrpArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerIysAsp | 94 |
| Db | 2322 | GGATGGAGAGGAGCAGGCTGTGACCTGGACCATGGAGACTCTTTTCACAGATAGCAAGGAC | 2381 |
| Qy | 95 | AsnAspGlyAspGlyLeuVal 101 | |
| Db | 2382 | AATGAAGGGATGGACTCATTT 2402 | |

RESULT 8
AX662355 8645 bp DNA linear PAT 22-MAR-2003

LOCUS Sequence 37 from Patent WO02062999.

DEFINITION AX662355

ACCESSION AX662355.1 GI:29163217

VERSION

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X.,
Malyankar,U., Shmkets,R.A., Tchernev,V.T., Spaderna,S.K.,
Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A.,
Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F.,
Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G.,
Millet,I. and MacDougall,J.R.

TITLE Proteins and nucleic acids encoding same

JOURNAL Patent: WO 02062999-A 37 15-AUG-2002;
Curagen Corporation (US)

FEATURES
source
1. .8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 3.1e-34 Length: 8645


```

RESULT 11
AKI25869
LOCUS      3394 bp      mRNA      linear      PRI 09-SEP-2003
DEFINITION Homo sapiens cDNA FLJ43881 fis, clone TEST14009028, highly similar
            to Mus musculus odd Oz/ten-m homolog 3 (Drosophila) (Odz3).
ACCESSION  AKI25869.1  GI:34532128
VERSION     oligo capping; fis (full insert sequence).
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
            Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,K.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Watanabe,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
            Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
            Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
            NEDO human cDNA sequencing project
            Unpublished
            2 (bases 1 to 3394)
            Isogai,T. and Yamamoto,J.
            Direct Submission
            Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
            NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
            Location/Qualifiers
            1. 3394
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="TEST14009028"
               /tissue_type="testis"
               /clone_lib="TEST14"
               /note="cloning vector: pME18SFL3"
FEATURES             source
     source           Location/Qualifiers
     1. 3394
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="TEST14009028"
        /tissue_type="testis"
        /clone_lib="TEST14"
        /note="cloning vector: pME18SFL3"
ORIGIN
Alignment Scores:
Pred. No.:      6.2e-32      Length:      3394
Score:          483.50      Matches:    78
Percent Similarity: 82.18%      Conservative: 5
Best Local Similarity: 77.23%      Mismatches: 9
Query Match:     78.36%      Indels:    9
DB:              Gaps:      1
US-10-029-020-14_COPY_750_850 (1-101) x AKI25869 (1-3394)
QY      1  TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db      491 TGGACGGCGCCAGCGCTGTAATCAGAGAGCGCTGCCACCCCGCTGTGCCGAGCAGCGGACC 550
QY      21  CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db      551 TGCAGAGTGGCAAGTGTGAATCAGCGCGCTGGAATGAGAGCACTGCACCTATC--- 607
QY      41  HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db      608 -----GAGGTTGTCTCTGTCTGTCTGTCACAGCAATGGAAGA 643
QY      61  CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db      644 TGTACCTGGACCAAAATGGCTGGCAATTGTGTGTGTCACCGCTGGATGGAGAGGACGAGC 703

```

```

81  CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyLeu 100
Db      704 TGTGACGTAGCATGAGACTCTTTGCCACAGATACAGGACATGAGGAGATGAGCTC 763
QY      101 Val 101
Db      764 ATT 766
RESULT 12
AF195418
LOCUS      7816 bp      mRNA      linear      ROD 27-JAN-2000
DEFINITION Mus musculus ODZ3 (Odz3) mRNA, partial cds.
ACCESSION  AF195418
VERSION     AF195418.1  GI:6760368
KEYWORDS    Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM    Mus musculus
REFERENCE   1
AUTHORS     Ben-Zur,T., Feige,E., Motro,B. and Wides,R.
            The mammalian Odz gene family: Homologs of a Drosophila pair rule
            gene with expression implying distinct yet overlapping
            developmental roles
            Dev. Biol. 117, 107-120 (2000)
            2 (bases 1 to 7816)
            Ben-Zur,T., Motro,B. and Wides,R.
            Direct Submission
            Submitted (15-OCT-1999) Faculty of Life Sciences, Bar-Ilan
            University, Ramat-Gan, Israel
            Location/Qualifiers
            1. 7816
               /organism="Mus musculus"
               /mol_type="mRNA"
               /db_xref="taxon:10090"
               <1. 7816
               /gene="Odz3"
               /note="synonym: Ten-m3"
               <1. 7042
               /gene="Odz3"
               /note="similar to the Drosophila melanogaster pair rule
               odd Oz/Ten-m gene sequence"
               /codon_start=2
               /product="ODZ3"
               /protein_id="AAP28316.1"
               /db_xref="GI:6760369"
               /translation="NSDTVTNTVSLPSGDNKLGKGGFTHTNTIDSGELDIGRAIOE
               VPPGIFWRSQLFIDQFLKFNLSLQKDALIGYGRKGLPPSHQTQYDFVELLDSRLI
               AREQRNLVESRAGRQARSVSHEAGFIQYDSGIWHLAFINDGKNPEQVSFTVIE
               SVVECPNCHNGECVSGTCHFPGLPDCSRAACFVLCSGNGQYSKGRCLCFSGMK
               GTEDVPTTCIDPQCGRGICIMSGACNYSKNGENECEDCLDFGCSNHCIGHE
               CHNCPMGSGNCEILATMCADQCSGHGTLYOESGCTCDPNMTGPDCSNHCISVDCCS
               HGVCMGSCRCCEGWTGPACNORACHPRCAHGTCKDKGCEKQWNGEHCITGCGPG
               LCNNGRCITLDQNGHCVCQCPQWRGAGCVAMETLTCTDSKDNESDGLIDCMDPCCLQ
               SLASVIRGQVLTADGTLGVNVSLFHYSEYDITITQDGMFDLVANGASLTLVFER
               SPFLTQVTVIWPNVFYMDTLVMKEENDIPSCDLSGFVRESPIVSSPLSTFRS
               SPDSPLIPETQVLHEETTPGTDLKLSYSSRAAGYKSLKITMTQAVIPFNLKVVH
               LMVAVGRLFOKMPFAPSNLAYTFINDTAYNQVYGLSEAVSYGYEYESCLDLTL
               WEKRTAVLQGTFLDASNGWTLDKHVLQVNGILYKNGENQFI SQOPPVYSIMG
               NNGRRSISPCSNQADKNLAPVALACGIDGSLYVDGPNYVRIPPSGNVTSVLEL
               SSNPARYLTADPTVTDGLYSDNTNRIYRPKSLTGAKOLTKNAEVVAGTGGCCLFP
               DEARCGDGKXAVEATLMSPKGMALDKNGLIYFVDMIRKVDQNGIISTLLGSDNLF
               ARPLTCDTSMHISOVLEWPTDLAINPMDNSIYLDNNVYLQITENRQVRIAGRPMH
               COVPEYVPVGHAVOTLESATAIYASVSGVLYITETDEKINRIQATDGTGIRSLV
               AGIPSCDCKNDACVQSGDGVAKDKNAPSLAASPDGTLIYIADLGNIRAVS
               XNPELLSNMFIYVASTPDQELYIFDINGHTQYVSLVTDGLYNFYSNDNDTAVT
               DSNGLNTRIRDPNRPVPSVDPNVFTTFTGNGCLKMTAQGLBLVLTFTHNSG
               LLATKSDTGTMTFFDYSRGLNVTNVTFTGVTNLHGDMDKAITLIESSRREDVS
               ITSNLSSIDSFTYTMVQDLRNSYQIGDGLRIFYASGLDSHYTETPHVLGAGTANPTV
               AKRNMTPGNGQNLVEMRFRKQAGKQVGRKLVNGNLLSVDFDPTTKTEKIY
               DDHRKFLRLIAYDTSGHFTLWLPSSKMAVNVVTSSTQIASIQRTTSEKVDVDSQ

```

RIVSRVADGKWTSTYILEKSMVLLHSHQRYIFEDMDRLSALTMPSEVAPHTMOTI
RSIGYRNLNPPESNASITIDYNEEGILLQAFIGTSSRVLPKPKRQTRLSIELYDS
TRVSTYDTEAGLVATNLOSDFCTIRYRQIGLDRQIFRFDSEGNWNAFYDYS
DNSRVTSMQGVNTEPTIDIDYQDDISGKQFGKFGVYIDINQIIISTAVMTYTK
HFDAGRIKEIYFRSLMWTITQYDNMGRTKREIKIGFPAFTTKYAYEYDVG
LQTVLNEKIMRWYNLNGNLNPPSSARLPLRDLRLDRITRLGDVOYVLDG
FLQRGTIFEYSSKGLTRVYSGKSGWTVIYRDLGRVSSKTSLSGHLQFFVADL
TYPTRI THVYNHSSSEITSLVYDLOGLHFLAMEISSGDEFYIASDNTGTPLVFSSNGL
MKLOIQTAYGEIYEDSNVDQLVGFPGGLYDPTKLHFGERYDIDLAGRWITPDI
ELWKIGDPAFNLPMFNPNPASKIDHVDKDIYDVSNLVTFGHLNNAIPGPVP
KFDLPESVYKVSQWEDVPIIFGVQOQVARQAFLSLGMAVQVSRKAGAEQS
WLWFAVKSLLGKVMYAVSQRVQTNVINTANEDCIKAAVLNNAFLNHLFTIEG
KDFHTYFKTTPTRLSRGLKALENGINTVSTTWNQRTREFADVMQEG
ALALHVRVGMILDEKARILBOARALARAWAREQOVRDGESEARLWTEGKRQLL
SAGKVGQYDGYVLSVEQYPELADSANNIQLRQSEIGKR

ORIGIN

Alignment Scores:
Pred. No.: 1.37e-31 Length: 7816
Score: 483.50 Matches: 78
Percent Similarity: 82.18% Conservative: 5
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 78.36% Indels: 9
DB: 10 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x AF195418 (1-7816)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 1046 TGGACCGCGCGCGTGAATCAGAGAGCTTGCACCCCTGCTGTGAGCAGCGGACG 1105
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 1106 TGCAAGCAGCGCAGTGGAGTGCAGCAAGGATGGACGAGAGCCTGCACAAATT--- 1162
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 1163 -----GAGGTTGCCCCGCTGTGCAACAGCAATGGGAGA 1198
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 1199 TGACACTGGACCAAAACGGCTGCACCTTTGCCAGCGCGTGGAGAGGAGCAGGC 1258
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 1259 TGTGACGTAGCCATGGAGACCTCTGTACAGACAGCAACAGGAGAGGAGCAGGACTC 1318
QY 101 Val 101
Db 1319 ATT 1321

RESULT 13
AX662359
LOCUS AX662359 8487 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 41 from Patent WO02062999.
ACCESSION AX662359
VERSION AX662359.1 GI:29163219
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
Malyanar, U., Shinkets, R.A., Tchernev, V.T., Spaderna, S.K.,
Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
Millett, I. and Macdougall, J.R.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 02062999-A 41 15-AUG-2002;
Curagen Corporation (US)
FEATURES Location/Qualifiers

source

1. .8487
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 3.94e-31 Length: 8487
Score: 478.50 Matches: 77
Percent Similarity: 81.1% Conservative: 5
Best Local Similarity: 76.24% Mismatches: 10
Query Match: 77.55% Indels: 9
DB: 6 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x AX662359 (1-8487)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2123 TGGACCGCGCGCGTGAATCAGAGAGCTTGCACCCCTGCTGTGCGCAGCAGCGGACC 2182
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2183 TGCAAGGATGCAAGTGTGAATGCAGCCATGGCTGGATGGAGAGCAGCATGCACTATC--- 2239
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2240 -----GAGGTTGCTGCTGTGCAACAGCAATGGAGA 2275
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2276 TGTACCTTGGACCAAAATGGCTGGCATGTGTGCGCAGCCTGGATGGAGAGGAGCAGGC 2335
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2336 TGTGACGTAGCCATGGAGACTCTTTGCACAGATAGCAGGACATGAGGAGATGGACTC 2395
QY 101 Val 101
Db 2396 ATT 2398

RESULT 14
AB025410

LOCUS AB025410 8373 bp mRNA linear ROD 08-MAY-1999
DEFINITION Mus musculus mRNA for Ten-ml, complete cds.
ACCESSION AB025410
VERSION AB025410.1 GI:4760775
KEYWORDS Ten-ml.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Ohashi, T., Zhou, X., Feng, X., Richter, B., Moergelin, M., Perez, M.T.,
Su, W., Chicquet-Ehrismann, R., Rauch, U. and Faessler, R.
TITLE Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
JOURNAL proteins expressed in many tissues
REFERENCE J. Cell Biol. (1999) In press
AUTHORS Ohashi, T.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
FEATURES Location/Qualifiers

source

1. .8373
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/dev_stage="adult"
/tissue_lib="brain"

GTISRODGSFDLVAIGGISVILIFDRSPFLPEKRTILWLPWNOPIVVEKVTMRVSD
PSCDISNFIPLSPPLSPFSGSCPERGTIVPELVQVEIPIPSSEFVRLSYLS
SRPGYKTLIRLLTHSTIPVGMKXHLTVAVEGRITQKMPFAAINLVITTFAMNKTDI
YQKVGIAEALVSVEYEYETCPDFILWEQTVVLQGFEMDASNLGDSWLNKHHILNP
QSGLIHKGNGENMFISQGPVISTIMNGHQRSVACTNCGPAHNKLFAPVALASGP
DGSVVGDFNFRRIFPSGNSVSLLELSTPAHKYILAMPVSESLYLSDTNTRKVKY
LKSIVETKDLKXNPEVAVAGTQDCLPDPQSHCGDGRASEASLNSPRGITVDRHGFY
FVDTGTMIRKIDENAVITTVIGSNGLTSTQPLSCDSCMDITQVRLWEMPTDLAVNPMDS
LYLVDNNILVQIGSENRRVLIAGRPICQVPGIDHFLVSKVAHSTLESARASVSHS
GLFIATDERKNRIQQVITINGEIIYIAGAPTDCKIDPNCDCFSGDGGYAKAKM
KAPSLAVSPDGLYVADLGNVRITIRNOAHLNMDNIYIAPADQELQYFTVNGT
HLHNLITRDYVNFYNSEGDGAIITSSNGNSVHIRRDAGGMLWLVVPGGVYWL
TISNGVLKRVSAQGVNPAIYTPGNTGLLATKSNENGWTVVEYDEGHLTNATPPT
GEVSSPHSDLEKLTKEVLDTSNRENVLMSTNLATSTIYILKOENTOSTYRVPDGS
RVTPASGMEIGLSEPHILAGAVNPTLGKCNISLPEHNANLEWRORKEQNKGNVSA
FERLRANHNLSIDFDHTRTKIYDDRHKFTLRILYDQTRPILWSPVSRVNEVN
ITYSPGLVTFIQRTWNEKXEDQSGKIIISRTWADGKIWSYTYLESKVMLLHSQRR
YIFEDQSDCLLSVTMPMSVRHSIQTMLSVGYRNIYTPDPSSTSFQDYSRDLIQ
TLHGTGRRLVLYKTKQARLSEVLYDTQVTLVYESSGVKTIHLMHDFICTIRYR
QTGLIGRQIFRSEEGLVNARPDYSVNNFRVTSQAVINETPLIDLYRYDVSGRT
EQPKFESVINYDLNQVITITVMKHTKIFPSANGQVIEQVIELKAIATWMTIQDNGR
HGNWCIKRVGDANIIRIFBYDADGQQLQTVSNDKIQWRYSYDINGDINLLSHGKSAR
LTPRLYDLRIITRLGIEIQYMDDEGELQQRGNDIFEYNSGGLLQKAYNKASGWTQY
YDGLGRVASKSLQHLQFFVDATNP IRVTHLYNHTSEITSLYDILQGHILAME
LSSGEYVACDNTGTPLAFVSRGOVKEILYTPYDGVHDTVPDFQVLIIGFHGGLY
DFTLKLHLGORDYDVVAGRWTTAYHHIKQLNLLPKPENLYSPENNYPVGKIQDVAK
YTTDILSWELEFGQLHNVLPGRPKPELENLELYELLRLQTKQEWDPKTIIGQOC
ELQKQANFSLSQLEPMTFRINDGRLEGQKPRFAVPSVFGKIGKFAIKDGLIVTAD
IIGVANDVSRLAAIENNAHYLENLHFTIEGRDTHYFIKLSLEEDILVIGNTGGRRI
LENGVNVTSQMTSLNGRTRRFADIQIOLGALCFNIRYGTIVTEEEKNHVLEIARORA
VAQAWTKQRRLQEGEIRAWTEGEKQQLSTGRVQGYGVLSVEQVLELSDSAN
NIHFMQSEIGRR"

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 6.02e-30 | Length: | 8297 |
| Score: | 464.50 | Matches: | 71 |
| Percent Similarity: | 82.00% | Conservative: | 11 |
| Best Local Similarity: | 71.00% | Mismatches: | 17 |
| Query Match: | 75.28% | Indels: | 1 |
| DB: | 9 | Gaps: | 1 |

US-10-029-020-14_COPY_750_850 (1-101) x AF100772 (1-8297)

| | | | |
|----|------|---|------|
| QY | 1 | TriMetGlyAlaAlaCysAspGlnArcAlaCysHisProArgCysAlaGluHisGlyThr | 20 |
| Db | 2210 | TGGGTAGGACCAACATGTGAGTGTAGCCCTGGATGGAGGGCCGACCACTGCACAAATGCT | 2269 |
| QY | 21 | CysArgAspGlyLysCysGluCysSerProGlyTyrAsnGlyGluHisCysThrIleAla | 40 |
| Db | 2270 | TGCAAGATGGAAAATCTGAGTGTAGCCCTGGATGGAGGGCCGACCACTGCACAAATGCT | 2329 |
| QY | 41 | HistTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg | 60 |
| Db | 2330 | CACACTTAGAT---GCTGTCCGAGATGGCTGCCAGGGCTCTGCTTGGAAATGGACGA | 2386 |
| QY | 61 | CysThrLeuAspLeuAsnGlyTyrHisCysValCysGlnLeuGlyTyrArgGlyAlaGly | 80 |
| Db | 2387 | TGTACCTGATCAAAATGGTTGGCACTGTGTGTGAGTGGGTGGAGTGGGACAGC | 2446 |
| QY | 81 | CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu | 100 |
| Db | 2447 | TGCAATGTTGTCATGGAAATGCTTTTGGAGATAACTTGGACAAATGATGGAGATGGTTTA | 2506 |

Search completed: August 14, 2004, 11:51:47
Job time : 1808.37 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 2670.71 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_1250_1400

Perfect score: 782

Sequence: 1 YIRIFPSGNVTNILELRNK.....LLGSNDLTARPLSCDSVMD 151

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US1002020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DS=genembl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZB=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1002020@cgn.1.19065 @runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pin.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|-----------|-------------|
| 1 | 782 | 100.0 | 8354 | 6 | AX556500 | Sequence |
| 2 | 767 | 98.1 | 8585 | 10 | AB025413 | Mus muscu |
| 3 | 759.5 | 97.1 | 8438 | 6 | AX675551 | Sequence |
| 4 | 727.5 | 93.0 | 8645 | 6 | AX600210 | Sequence |
| 5 | 712.5 | 91.1 | 9722 | 10 | AF059485 | Sequence |
| 6 | 625 | 79.9 | 8409 | 5 | GGA279031 | Mus muscu |
| 7 | 625 | 79.9 | 8409 | 5 | GGA279031 | Mus muscu |
| 8 | 619 | 79.2 | 3394 | 9 | AK125869 | Sequence |
| 9 | 619 | 79.2 | 3394 | 9 | AK125869 | Sequence |
| 10 | 619 | 79.2 | 5309 | 9 | AB040888 | Homo sapi |
| 11 | 619 | 79.2 | 8473 | 6 | AX662357 | Sequence |
| 12 | 619 | 79.2 | 8473 | 6 | AX662357 | Sequence |
| 13 | 619 | 79.2 | 8645 | 6 | AX662355 | Sequence |
| 14 | 618 | 79.0 | 8675 | 6 | AX662353 | Sequence |
| 15 | 618 | 79.0 | 5804 | 10 | AK122513 | Sequence |
| 16 | 613 | 78.4 | 6560 | 6 | AX250065 | Mus muscu |
| 17 | 613 | 78.4 | 7781 | 9 | AB032953 | Sequence |
| 18 | 613 | 78.4 | 9729 | 6 | AX250013 | Sequence |
| 19 | 613 | 78.4 | 9826 | 6 | AX250008 | Sequence |
| 20 | 610 | 78.0 | 8689 | 6 | AX250067 | Sequence |
| 21 | 610 | 78.0 | 8689 | 10 | AF086607 | Rattus no |
| 22 | 610 | 78.0 | 8797 | 6 | AX250063 | Sequence |
| 23 | 610 | 78.0 | 8797 | 6 | AX250066 | Sequence |
| 24 | 610 | 78.0 | 8797 | 10 | AB025411 | Mus muscu |
| 25 | 608.5 | 77.8 | 9264 | 5 | AB026980 | Danio rer |
| 26 | 602.5 | 77.0 | 8575 | 6 | AX921803 | Sequence |
| 27 | 588 | 75.2 | 8373 | 10 | AB025410 | Mus muscu |
| 28 | 585 | 74.8 | 8118 | 5 | GGA238613 | Gallus ga |
| 29 | 571.5 | 73.1 | 7713 | 9 | HSM808325 | Sequence |
| 30 | 563.5 | 72.1 | 7816 | 10 | AF195418 | Mus muscu |
| 31 | 555.5 | 71.0 | 8816 | 5 | AB026979 | Mus muscu |
| 32 | 534.5 | 68.4 | 8297 | 9 | AF100772 | Danio rer |
| 33 | 413 | 52.8 | 112872 | 2 | AP000716 | Homo sapi |
| 34 | 413 | 52.8 | 165812 | 2 | AP001141 | Homo sapi |
| 35 | 413 | 52.8 | 170156 | 2 | AP002412 | Homo sapi |
| 36 | 413 | 52.8 | 186084 | 9 | AP002768 | Homo sapi |
| 37 | 406 | 51.9 | 245722 | 2 | AC120288 | Homo sapi |
| 38 | 406 | 51.9 | 257693 | 2 | AC103323 | Rattus no |
| 39 | 405 | 51.8 | 226256 | 10 | AC073599 | Mus muscu |
| 40 | 388 | 49.6 | 179685 | 5 | BC001057 | Zebrafish |
| 41 | 347 | 44.4 | 321049 | 2 | AC137221 | Rattus no |
| 42 | 344 | 44.0 | 193544 | 5 | AL807754 | Zebrafish |
| 43 | 308.5 | 39.5 | 60511 | 9 | HS18181 | Human DNA |
| 44 | 307 | 39.3 | 146071 | 10 | AL929061 | Mouse DNA |
| 45 | 307 | 39.3 | 172465 | 2 | BX324192 | Mus muscu |

ALIGNMENTS

RESULT 1

```
AX556500
LOCUS AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Gangolli, E.A., Patturajan, M., Vernet, C.A., Malyankar, U.M.,
Kekuda, R., Stone, D.J., Anderson, D., Shinkens, R.A., Burgess, C.E.,
Zehrhusen, B.D., Liu, X., Spytek, K.A., Casman, S.J., Boldog, F.L.,
Smithson, G., Li, L. and Ji, W.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
FEATURES
source
1..8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4,41e-79 Length: 8354
Score: 782.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1250_1400 (1-151) x AX556500 (1-8354)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3782 TACATTTAGAGGATCTTCCCTCTGGAATGTCACCAATCTAGAGTCAGGAATAAA 3841
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrIleuAlaThrAspProMetSer 40
Db 3842 GATTTTCAGACATAGTCACAGTCACAGCACAAATACTACTCGGCCACAGACCCCATGAGT 3901
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3902 GGGCGCGCTTCTCTTCTGCACACAGCGCGCGGCTCTTTAAATCAAGTCCATGTCG 3961
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3962 GTGCTGAAGACCTTGTCAAGAACTCTGAGTGTGTTGGGGGACAGTGACCACTGCCTC 4021
QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4022 CCCTTTGATGACACTCGCTCGGGGATGTTGGGAAGCCACAGAACCCACACCAAT 4081
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4082 CCAGGGGATATTACAGTGGCAAGTTTGGCTCTGATCTACTCTGCGATGCGACCAATGATC 4141
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrIleuLeuGlySerAsnAspLeuThrSer 140
Db 4142 AGACGATCATCAGATGGATCATCTCCACCTCTGCTCGCTTAATGATCATCATCA 4201
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4202 GCCCGGCCACTGAGTGTGATCTTCATGATGAT 4234
RESULT 2
AB025413
LOCUS AB025413 8585 bp mRNA linear ROD 08-MAY-1999
DEFINITION Mus musculus mRNA for Ten-m4, complete cds.
ACCESSION AB025413
VERSION AB025413.1 GI:4760781
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Ohashi, T., Zhou, X., Feng, K., Richter, B., Moergelin, M., Perez, M.T.,
Su, W., Chiquet-Ehrismann, R., Rauch, U. and Faessler, R.
TITLE Mouse Ten-m/odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
JOURNAL J. Cell Biol. (1999) In press
REFERENCE
AUTHORS Ohashi, T.
Direct Submission
2 (bases 1 to 8585)
Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8552, Japan
(E-mail: ohashi@cc.okayama-u.ac.jp, Tel:+81-86-235-7128,
Fax:+81-86-222-7768)
FEATURES
Location/Qualifiers
1..8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1..8585
/gene="ten-m4"
188..8503
/gene="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"
/translation="MDVKKPKYRSLTRRDAERYVTSSADSEEGKPKQKSVSSSET
LKAYQDQDLAYSRVKMVPQEAEEFCRTGNFTRELGLGEMPPHCHTLVRTDGL
PCGYSGSSADSLADLTVLSPHPVRLWGRSTRSGSSCLSSRANSNLTLDTRDE
NTEITHPSLQNHPLRLTPPELPHANTPHQHAASINSLNRGFTFRNPPGPAFTDH
SLAGSPGAGAOEPHAQONWLLNSNIPLETNLGKQPPFLQDNLIMDLASRSH
DSYSDGHEFLPKGGTSLFCTPCYPLTSTVSPPLPRSTFSRPAFLNKKPS
KYCNKCAALSAILSATLIVLLAVFAMHLFGLNWLQPMEGOMQMYEITDASW
PYTDVLSVPSGGTGLTLETPDRKGAAGKSPSLPDSFDIDSGEIDVRRASQKTPP
GTFWRSQVFDHPVHLKFNVSLGKAALVGIYGRKGLPPSHTFDFVELLDGRLLTQE
ARSLGEPQSRGPPVPPSSHETFTYQDLSGIWHLAFVNDGKESVVSFLLTAIRSVD
NCPSCYNGQDCISGTHCFLGFLGDCRASCPVLCSNGKRCCLCHSGWKGA
CVPTNQCIDVACSSHGTCIMGTICINPGYKGSCEVDCMDPTCSRSGVCVRGECH
SVGWGTCETPRATCLDQSGHGTFLPDTGLCNDPSPWTGHDCEIICAADCGGHV
CVGTGCRBDGMMGAACDORACHPRCAEHGTCDGKCECSGWNGBHCTIAHLYDRV
KEGCPQLGNGRCITLDLNGHCVCOLGWRGTCGDTSMETGCGDGDNDGDGLVDCM
PNCPLQCHVNPCLIGSPDLDIIQETQAPVQNLNPFYDRIKELVGRDSTHSPG
ENPFDGHACVIRGQVMTSDTPLGVNIFINNPLFGYISRDGSDFLVINGGISI
ILRFERAPITQETHTLWLPDRFVFMETIMRHEENIEIPSCDLSNFARNPVVSPSL
TSFASCAKGPVPEIQALQEEIVLAGCKMRLSYLSRTPGYSKVLRLSLTHPTIPF
NMLHLMHVAVEGRLEFKWFAAPDLSYFIMDKTDVYQKVFGESEAFVSQYVEFS
CPDLIWEKRTAVLQGTIDASKIGWSLDKHALNIQSGILHKGNGQFVSQOPPV
IGSIMNGRRRSISCSNGLADGNKLPALVATCGSDGLYVGFYIRFIRFPGSNV
TNILEMRNDKFRHSHPAHKYLATDPMGAGFSLDTSNRRVFKVSTYVVKOLKNS
EVVAGTQDCLPFDTRCGDGKATEATLTPRGITVYDKGLIYFVDTGMIRVDQNG
IISTLGSNDLTSARPLSCDVMIEISQVLEWETDLAINPMDNSLVLDNNVLIQISE
NHQVRIVAGPCHQVPGIDHFLLSKVAIATLESATALASHNGVLYIATDEKIN
RIHQVTSIGSILVAGPSCDCKNDANCDSPGDGQYAKADKLNTPSSLAVACDGL
RIVAGTISGILVAGPSCDCKNDANCDSPGDGQYAKADKLNTPSSLAVACDGL
YVADLGNIRIRIRKKNRPLNTQNYELSPIDQELFLPDTSGKHLTQSLPFGDLY
NFTYTGDDITHITDNNNMNVNRDSTGMPLVVPDGVVYTPWTGINSALRSVTTQ
GHELMWYHNSGLLATKSNENGTTFYEDYFGRITNTVFTFGQVSSRSDTSV
HVOVETSSKDDVTITNLSAGAFYLLQDQVNSYIGADGSLRLANGREVALQT
EPHLLAGTVNPTVTKENVTLPIDNGLNLEWRQKQEQARQVTVFGRRLVRHNLIS
LDPRVTRTEKIYDHRKFLRLILYDQAGPSLWSPSRRLNGVNTYSPGHIAIOR
GIMSERMEYDQAGRTISRIFADGKMSYTYLEKSMVLHLHSQRYIFEFKNDRLSSV
TMPNFARQTLTIRSVGYRNIYQPPPEGNASVLDQFTEGDHLLHTFVLTGRRVLYKY
GKLSKLAETLYDTTFSYVDETAGMLKTNLQNEGFTCTIRYQIQGLDQLRFRFT
EGMVMARFNDYDNGSPRVTSMQAVINETPLDLYRIDYDVSQGTGFGKPGVLYDI
NQIITTAVMTHTKHFDAYGRMKVQYEIFERSLWYMTVQYDNNMGRVVKELKGVFPAN
```

TTTYSVEYDADGQLOTVSINDKPLWRYSDYDNGNHLHSPGNSARLTPLRYDLDRIT
RLGDVQYKMDDEGFLRQGDVFEYNSAGLLIKAYNRASGWSVRVYDGLGRVSXS
SHSHLOFVADYDNTPTKVTYHNSSEITSLYDLOHLFAMELSSGDFEYVACDN
ICTPLAFVSTGLMLKQLLYTAYGEIYMDTNEFQIIIGHGLYDPLTKLVHMGRRD
YDVLARGSPHEMLWKRUSNSIVFPHLYMKNPNFINSQDIKCFQVQKQKAFVTL
GFQHNVPYGPDPDMEPSYELVMTQKTDWNSKSIILGVQCEVQKQKAFVTL
ESFDQYSGTITSCQAPETKFPASGSIKGVKVPALKDGRVTTDIIISVANEDGRRI
AAILNHAHYLENLHFTIDGVDTYHFKPGPSGLDAILGLSGRRTLENGVNTVSOI
NTMLSGRTRYTDIOLOYRACLNTRYGTTVDDEKVRVLELQARQAVRQAWARQORL
RGEBSGLRATWDEKQOVLNTRGVQYDGFVTSVEQYFELSDSANNHFMQSEMR
R"

ORIGIN

Alignment Scores:
Pred. No.: 2,41e-77 Length: 8585
Score: 767.00 Matches: 145
Percent Similarity: 99.34% Conservative: 5
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 98.08% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AB025413 (1-8585)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3941 TACATCCCGAGAATCTTCCCTCTGGAATGTCACCAACATCTCGAGATGAGAAATAA 4000
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 4001 GATTTACAGACATAGTACAGCCCGACACACAAATACTACCTGGCTACAGCCCCATGAGT 4060
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLeuIleLysSerThrVal 60
Db 4061 GGGGCGGCTTCTCTGTCTGACACCAACAGCCCGCGGGTCTTCAAGTCAAGTCCACCA 4120
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 4121 GTGGTGAGGACCTGGTGAAGAACTCCGAGGTGGTAGCAGGAGCTGGTACAGTGCCTC 4180
QY 81 ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4181 CCCCTTGTATGATACCCGCTGGAGATGGTGGGAAAGCCACAGAGCCACGCTCACTAAC 4240
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetile 120
Db 4241 CCCAGGGAAATTACAGTGGCAAGATTTGGGCTCATTTATTTCGGAGCGGACCATGATC 4300
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4301 AGAGTGTGATCAAAATGGAATCATCTCCACTTTGCTGGGCTCCAATGACCTCACCTCG 4360
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4361 GCCAGGCCCTCAGCTGTGACTCCCTCATGGAG 4393

RESULT 3
AX675551
LOCUS AX675551 8438 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 1 from Patent WO02055704.
ACCESSION AX675551
VERSION AX675551.1 GI:29333552
KEYWORDS Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Padigar,M., Li,L., Zerhusen,B.D., Casman,S.J., Shenoy,S.,
Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M.,
Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E., Guo,X.,
Baldog,F.L., Grosse,W.M., Alsobrook,J.P., Gerlach,V.,
Baidermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J.,
Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and

Stone,D.J.
Proteins, polynucleotides encoding them and methods of using the
same
JOURNAL Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source
1..8438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1.72e-76 Length: 8438
Score: 759.50 Matches: 151
Percent Similarity: 96.79% Conservative: 0
Best Local Similarity: 96.79% Mismatches: 0
Query Match: 97.12% Indels: 5
DB: 6 Gaps: 2

US-10-029-020-14_COPY_1250_1400 (1-151) x AX675551 (1-8438)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeu-----Arg 18
Db 3764 TACATTAGAAGGAUUTTCCTCTGGAATGTCACCAACATCTCTAGAGCTGAGGTCAGA 3823
QY 19 AsnLysAspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPro 38
Db 3824 AATAAAGATTTTCAGACATAGTCAGATCCAGCACACAAATACTACTCTGCCACACAGCCCC 3883
QY 39 MetSerGlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSer 58
Db 3884 ATGAGTGGGGCGGCTTCTCTTCTGACAGCAACAGCCGGCGGTCTTTAAAAATCAAGTCC 3943
QY 59 ThrValValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGln 78
Db 3944 ACTGTGGTGGTGAAGACCTTGTCAAGAACTCTCAGGTGGTGGGGGACAGGTACACAG 4003
QY 79 CysLeuProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeu 98
Db 4004 TGCCTCCCTTTGATGACATCGCTCGGGGATGGTGGGAGGCCACAGAGCCACACATC 4063
QY 99 ThrAsnProArg-----GlyIleThrValAspLysPheGlyLeuIleTyrPheVal 115
Db 4064 ACCAATCCCGAGGGTCCCGCAGGCAITACAGTGGCAAGATTTGGGCTGATCTCTCGTG 4123
QY 116 AspGlyThrMetileArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySer 135
Db 4124 GATGGCACCATGATCAGACGCAATCGATCAGATGGGATCATCTCCACCTGCTCGGTCT 4183
QY 136 AsnAspLeuThrSerAlaArgProLeuSerCysAspSerValMetAsp 151
Db 4184 AATGATCTCATCATCAGCCCGCCACTCAGCTGTGATTCTGTCTGATGAT 4231

RESULT 4
AX600210
LOCUS AX600210 8645 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S.,
Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,
Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K.,
Khare,R. and Wallia,N.K.
Proteins associated with cell growth, differentiation, and death
Patent: WO 02072830-A 22 19-SEP-2002;
Incyte Genomics, Inc. (US)

FEATURES
source

Location/Qualifiers
1. .8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7498573CB1"

Location/Qualifiers
1. .9722
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/cell_line="NIH-3T3"
/note="tunicamycin-treated"

ORIGIN

Alignment Scores:
Pred. No.: 8,43e-73 Length: 8645
Score: 727.50 Matches: 144
Percent Similarity: 95.3% Conservative: 0
Best Local Similarity: 95.3% Mismatches: 0
Query Match: 93.0% Indels: 7
DB: 6 Gaps: 1

gene

/gene="Doc4"
/note="downstream of CHOP 4"

CDS

583..9060
/gene="Doc4"
/note="similar to Drosophila melanogaster temm/odz and human gamma-hergulin; type II transmembrane protein"

misc_feature

/product="DOC4"
/protein_id="AAC31807.1"
/db_xref="GI:3170615"

ORIGIN

US-10-029-020-14_COPY_1250_1400 (1-151) x AX600210 (1-8645)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3852 TACATTAGAAGATCTTCCCTCTGGAAATGTCACCAATCTTAGAGCTG----- 3902
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3903 -----AGTCACAGTCCAGCACACAATACTACTCTGGCCACAGACCCCATGAGT 3950
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3951 GGGGCGGTCTCTCTTCTCACAGCAACAGCGCGGGTCTTTAAATCAAGTCCATCTGTG 4010
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 4011 GTGGTGAAGACCTTCTCAAGAACTCTGAGGGTGTGGGAGGACAGTACACAGTCCCTC 4070
QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyValAlaThrGluAlaThrLeuThrAsn 100
Db 4071 CCCTTTGATGACACTCGCTCGGGGATGTGGGAAGCCACAGACCACTACCAAT 4130
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4131 CCCAGGGCATTACAGTGACAAAGTTGGGCTGATCTACTTCGTGGATGACCATGATC 4190
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4191 AGACGATGATCAGATGGATGGATCTCCACCTGCTCGGCTTAATGATCTCATCATCA 4250
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4251 GCCCGGCCACTCAGCTGTGATTCGTGATGAT 4283

misc_feature

1813..1881
/gene="Doc4"

ORIGIN

/note="putative; transmembrane-region site"

RESULT 5
AF059485
LOCUS Mus musculus Doc4 (Doc4) mRNA, complete cds. 9722 bp linear ROD 15-AUG-1998
DEFINITION Mus musculus Doc4 (Doc4) mRNA, complete cds.
ACCESSION AF059485
VERSION AF059485.1 GI:3170614
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE 1 (bases 1 to 9722)
AUTHORS Wang,X.-Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P., Zinsner,H. and Ron,D.
TITLE Identification of novel stress-induced genes downstream of chop
JOURNAL EMBO J. 17 (13), 3619-3630 (1998)
MEDLINE 98315054
PUBMED 949432
REFERENCE 2 (bases 1 to 9722)
AUTHORS Wang,X.-Z. and Ron,D.
TITLE Direct Submission
JOURNAL Submitted (14-APR-1998) Skirball Institute, New York University Medical Center, 550 First Ave., New York, NY 10016, USA

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 5,14e-71 | Length: | 9722 |
| Score: | 712.50 | Matches: | 138 |
| Percent Similarity: | 94.70% | Conservative: | 5 |
| Best Local Similarity: | 91.39% | Mismatches: | 1 |
| Query Match: | 91.11% | Indels: | 7 |
| DB: | 10 | Gaps: | 1 |

US-10-029-020-14_COPY_1250_1400 (1-151) x AF059495 (1-9722)

| | | | |
|----|------|--|------|
| QY | 1 | TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys | 20 |
| Db | 4519 | TACATCCGAGAAATCTCCCTCTCGAAATGTCACCAATCTCTGAGATG----- | 4569 |
| QY | 21 | AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer | 40 |
| Db | 4570 | -----AGTCACAGCCCAACACCAAACTACTACCTGGCTACAGACCCCATGAGT | 4617 |
| QY | 41 | GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal | 60 |
| Db | 4618 | GGGGCCGCTCTCTCTGTGACACCAACAGCCGGCGGTCTTCAAGTCAAGTCCACCA | 4677 |
| QY | 61 | ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu | 80 |
| Db | 4678 | GTGGTGAAGGACCTGGTGAAGAACTCCGAGGTGGTAGCAGGACTGGTACCACTA | 4737 |
| QY | 81 | ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn | 100 |
| Db | 4738 | CCCTTTCATGATACCCGCTGGGAGATGGTGGGAAGCCACAGAAGCCACGCTCA | 4797 |
| QY | 101 | ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle | 120 |
| Db | 4798 | CCACAGGGAATACAGTGGCAAGTTTGGGCTCATTTATTCGTGGAGCGACCATG | 4857 |
| QY | 121 | ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer | 140 |
| Db | 4858 | AGACGTGTTGATCAAAATGGAATCATCTCCATTTGCTGGGCTCCAATGACCTCA | 4917 |
| QY | 141 | AlaArgProLeuSerCysAspSerValMetAsp | 151 |
| Db | 4918 | GCCAGGCCCTCAGCTGTGATCTCCGTGATGGAG | 4950 |

RESULT 6
GGA279031
LOCUS
DEFINITION
Gallus gallus partial mRNA for teneurin-2 (TEN2 gene), long splice variant.
ACCESSION
AJ279031
VERSION
AJ279031.1 GI:10241573
KEYWORDS
ten2 gene; teneurin-2.
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE
1
Tucker,R.P., Chiquet-Ehrismann,R., Chevrone,M.P., Martin,D., Hall,R.J., and Rubin,B.P.
Teneurin-2 is expressed in tissues that regulate limb and somite pattern formation and is induced in vitro and in situ by FGF8
Dev. Dyn. 220 (1), 27-39 (2001)
20581705
PUBMED
11146505
REFERENCE
2 (bases 1 to 8409)
Chiquet-Ehrismann,R.
Direct Submission
AUTHORS
Submitted (13-SEP-2000) Chiquet-Ehrismann R., Cell Biology, Friedrich Miescher Institute, Maulbeerstrasse 66, CH-4058 Basel, SWITZERLAND

FEATURES
source
1..8409
/organism="Gallus gallus"
/mol_type="mRNA"

Location/Qualifiers

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:

4.9e-61
625.00
90.07%
74.83%

Length:
Matches:
Conservative:
Mismatches:

8409
113
23
15

mat_peptide

<1..>8406

/gene="TEN2"

/product="teneurin-2"

/evidence=experimental

/db_xref="taxon:9031"
/tissue lib="11 day and 14 day embryonic brain; 11 day embryonic retina"
1..8409
/gene="TEN2"
<1..>8406
/gene="TEN2"
/function="pattern formation and neuronal development"
/codon_start=1
/evidence=experimental
/product="teneurin-2"
/protein_id="CAC09416.1"
/db_xref="GI:10241574"

gene
CDS

| | | | |
|--|---|--|----------------------------|
| Query Match: | 79.92% | Indels: | 0 |
| DB: | 5 | Gaps: | 0 |
| US-10-029-020-14_COPY_1250_1400 (1-151) x GGA279031 (1-8409) | | | |
| QY | 1 | TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys | 20 |
| DB | 3856 | TACATTCGGCGTATCTTCCCATCCAGGAATGTGACTAGCATATTGGAGCTGAGAAATAAA | 3915 |
| QY | 21 | AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer | 40 |
| DB | 3916 | GAGTTTAAACATAGCAACAATCTCTGCTCACAATACTATCTGGCGGTGGACCCCGTTTCG | 3975 |
| QY | 41 | GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal | 60 |
| DB | 3976 | GGCTCCCTGTACGTATCAGACCAACAGCGGATATCAAAAGTCAAAATCTTCTACT | 4035 |
| QY | 61 | ValValLysAspLeuValLysAsnSerGluValAlaGlyThrGlyAspGlnCysLeu | 80 |
| DB | 4036 | GGCACAAGACCTGGCTGTAATCTGAAGTGGTAGCGGGACTGGAGAGCAATGCGCTG | 4095 |
| QY | 81 | ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn | 100 |
| DB | 4096 | CCCTTTGATGAGCCAGATGTGGAGATGGAGGAAAGCAGTGGAGCAACCTAATGAGT | 4155 |
| QY | 101 | ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle | 120 |
| DB | 4156 | CCTCGAGGAATTCGAGTGGATAGTATGACTCATGTATTTTGTGTAGCCACTATGATT | 4215 |
| QY | 121 | ArgArgIleAspGlnAsnGlyIleLeuSerThrLeuLeuGlySerAsnAspLeuThrSer | 140 |
| DB | 4216 | CGAAAGTGGATCAGAAATGGAATATATCACTCTGCTGGGCTCCAATGACCTAATGCC | 4275 |
| QY | 141 | AlaArgProLeuSerCysAspSerValMetAsp | 151 |
| DB | 4276 | GTCCGACCTCTAAGCTGTGATTCACGATGGAT | 4308 |
| RESULT 7 | | | |
| LOCUS | AX250068 | 8409 bp | DNA linear PAT 28-SEP-2001 |
| DEFINITION | Sequence 67 from Patent WO0166747. | | |
| ACCESSION | AX250068 | | |
| VERSION | AX250068.1 GI:15864504 | | |
| KEYWORDS | Gallus gallus (chicken) | | |
| ORGANISM | Gallus gallus | | |
| REFERENCE | 1 Vernet, C.A., Fernandes, E., Shinkets, R.A., Herrmann, J.L., Majumder, K., Macdougall, J., Mishra, V., Mezes, P.S. and Rastelli, L. Proteins named fctxr and nucleic acids encoding same Patent: WO 0166747-A 67 13-SEP-2001; | | |
| AUTHORS | Curagen Corporation (US) | | |
| TITLE | Location/Qualifiers | | |
| JOURNAL | 1. 8409 | | |
| FEATURES | /organism="Gallus gallus" | | |
| source | /mol_type="unassigned DNA" | | |
| | /db_xref="taxon:9031" | | |
| ORIGIN | Alignment Scores: | | |
| | Pred. No.: | 4.9e-61 | Length: 8409 |
| | Score: | 625.00 | Matches: 113 |
| | Percent Similarity: | 90.07% | Conservative: 23 |
| | Best Local Similarity: | 74.83% | Mismatches: 15 |
| | Query Match: | 79.92% | Indels: 0 |
| DB: | Gaps: 0 | | |
| US-10-029-020-14_COPY_1250_1400 (1-151) x AX250068 (1-8409) | | | |
| QY | 1 | TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys | 20 |


```
/clone="TESTI4009028"
/tissue_type="testis"
/clone_lib="TESTI4"
/note="Cloning vector: pME18SF13"

ORIGIN
Alignment Scores:
Pred. No.: 7,99e-61 Length: 3394
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservativity: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AK125869 (1-3394)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 1964 TACGTGGCGCGGATATTCCTTCTGAAATGTAACAAGTCTTAGAATTAAGAAATAA 2023
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 2024 GATTTTAGACATAGCAGCAACCCAGCTCATGATACCTTGCACACCGATCCAGTCACG 2083
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrVal 60
Db 2084 GGAGATCTGTACGTTCTTGACACAAACACCCGAGATTTATGCCCAAGTCACTTACG 2143
QY 61 ValValIysAspLeuValIysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 2144 GGGCAAAAGACTTGACTTAAATAATGCAGAGTCGTCGAGGAGCGGAGCAATGCCTT 2203
QY 81 ProPheAspThrArgCysGlyAspGlyValGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 2204 CCGTTTGACGAGCGCAGATGCGGATGAGGAGGAGCGCGGAGGAGCCACATCATGAGT 2263
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 2264 CCAAAAGAAATGCGAGTTGATAAGAAATGATTAATCTACTTTGTTGATGAAACCATGATT 2323
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 2324 AGAAAGTTGACCAAAATGGAATCATATCAATCACTCTTCTGGCTCTACGATTTGACTTCA 2383
QY 141 AlaArgProLeuSerCysAspSerValMet 150
Db 2384 GCCAGACCTTAACTTGTGTGACACAGCATG 2413

RESULT 9
AB040888
LOCUS Homo sapiens mRNA for KIAA1455 protein, partial cds.
DEFINITION Homo sapiens mRNA for KIAA1455 protein, partial cds.
ACCESSION AB040888
VERSION AB040888.1 GI:7959170
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
Nagase,T., Kikuno,R., Ishikawa,K., Hirose,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVII. The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro
DNA Res. 7 (2), 143-150 (2000)
JOURNAL 20277482
MEDLINE 10819331
PUBMED
REFERENCE 2 (bases 1 to 5309)
Chara,O., Nagase,T. and Kikuno,R.
Direct Submission
AUTHORS
TITLE Submitted (04-APR-2000) Osamu Chara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnaiffo@kazusa.or.jp)
```

URL: <http://www.kazusa.or.jp/huge>, Tel: 81-438-52-3913,
Fax: 81-438-52-3914

FEATURES
Location/Qualifiers
1..5309

source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fh16070"
/clone_lib="pBluescriptII SK plus"
1..5309
/gene="KIAA1455"
<1..>5309
/gene="KIAA1455"
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1455 protein"
/protein_id="BAA95979.1"
/db_xref="GI:7959171"
/translation="QNPQYCRGLPDDPDIISQLSQSPSQQAQKSFYDRISFLIGSDST
HVIPEGSPFNKSLASVIRGQVLTADGTPGIVNVSFFHYPEYGVITIQDGMGFDLVA
GGASLTIVFERSPLTQVHTVIMVFNVMYMDTLVKKRENDIPSCOLSQVFRNPIL
VSSPLSTPFRSSPEDSPILPETQVLHETIIPGDTLKLKLSRAAGTKSVKLTMTQ
SLIPNLKVLHVAVVGRLPQKVPASPNIATYFIWKTDAYNQKVYGLSEAVSVSG
YEYSCDLTILWEAKTALIQYELDASNMGGWTLDKHVLVDQNGILYKGNQFIS
QQPVVSSIMNGRRRSISPCSCNQADGNKLAPALACGIDGSLYGVDFNVRRI
PSGVTSLVLELRNKRDFRHSNPAHRYLATDPTVGDLYSDTNTRRIVRPKSLTGAKD
LTKNAEVAGTGEOLPDEARPCGCGKAVEATLSPKGMADVKNGLIYFVDTMIRK
VDQNGIISTLLGSNDLTSARPLCTDTSMLISQVLEWFTDLAINPMONSITVLDNNVV
LQITENQVRLAAGRMHCQVGPVGPVGHAVQTTLLESATAIAVSYGVLIITETDE
KLNIRKQVTDGELISLVAGIPSECDCKNDANCDYQSGDGYAKDAKLSPSSLAASP
DGLTIADLGNIRIRAVSKNKPILLNSMNFYEASPTDQELIFIDINGTHQVTVSLVG
DLYNFYSNDNDITAVTNGNLTLRIRRDNPMPVRVSPNOVIMLTIGTNCGLK
MTAGLELVLTVHGNSGLATSDGTGTTFFDYDSEGLTNVTPTGVVYNLAGDM
DKALTVDIESSREEDVSIITSNLSSIDSFYTMQDQLRNSYQIGYDGLRIIYASGLD
SHYQTEPHVLACTANPFAKXNMTLPGENQNLVEMRFRKQAKGVNVFGRKVRNG
RNLLSDVDFRTIKIYDHRKFLRIAIDTSGHPTLWLPSSKLMAVNVYTSGTQI
ASLQRTTSEKVDYDQGRIVSRVFDAGKTSYILEKSMVLLHSQRQVITFYDMD
RLSALTMPVSAREMTQIRSYRNINYPNESNASITIDYNEEGLLQITVAFGLTSRR
LFPKFRQTLNARFIDYDSRVSTYDTAGVAKTVNLQSDGFICTIRVOIGPLIDRQ
IPFSDGMVNRARFSDYNSFRVTSMOGVINETPLPIDLYQFDDISGVEQFGKFGV
IYDINQIISTAVMTYTKFHDAHRIKEIQEYIFRSLMWITITQYDNMGRVTKREIK
GRPATIKAYEYDQGVQTVYINERKIMMYNDLNGNLHLLNPSARILTPRYDL
RRTIKLGDVQVRLDDEGLRQGTETFEYSSKGLLTVYSGKSGVTIVRYDGLGR
VSSKSLGHLQFPYADLTPTITHVYHNSSEITSLYDILQHLFAMEISGSDRH
IASDNTGTPLAFFSSNGLMKQIYATGEIYFDSNIDFOLVIGHGLYDPLTKLH
FGERDYDILAGWTTPTDIEIKRIGKDPAPFNLYMERNNNPASKIHVDKDIYDVSW
LVTDFPHLHNAIPGFVPKFDLTSPSYELVKSQWDDIPFVGQQQVAKAFASL
GKMAEVQSVRRAGGASWLPATVATKSLIGKVMIAVSQGRVQTNVINTIANEDCIKVA
AVLNNAFYLENLHFTIEGDKTHYFIKTTPTESDLGLRLTSGR"

ORIGIN

Alignment Scores:
Pred. No.: 1,37e-60 Length: 5309
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservativity: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AB040888 (1-5309)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 1161 TACGTGGCGCGGATATTCCTTCTGAAATGTAACAAGTCTTAGAATTAAGAAATAA 1220
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 1221 GATTTTAGACATAGCAGCAACCCAGCTCATGATACCTTGCACACCGATCCAGTCACG 1280
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrVal 60
Db 1281 GGAGATCTGTAGTCTTCTGACACAAACACCCGAGATTTATGCCCAAGTCACTTACG 1340

Mon Aug 16 09:01:05 2004

QY 61 ValVallysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
 Db 1341 GGGGCAAAAGACTTCACTAAAAATGCAAGATGCGGAGGACAGGGAGCAATGCCCTT 1400
 QY 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
 Db 1401 CCCTTTGACGAGGCGAGATGTGGGATGAGGGAAGCGCGTGGAGCCACCACTCATGAGT 1460
 QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetile 120
 Db 1461 CCCAAGCAATGGCAGTTGATAAGATGATTAATCTACTTTTGTGTGATGGAACCATGATT 1520
 QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 Db 1521 AGGAAGTTGACCAAAATGGAATCATATCACTCTCTGGGCTCTTAACGATTGTACTTCA 1580
 QY 141 AlaArgProLeuSerCysAspSerValMet 150
 Db 1581 GCCAGACCTTTAACTTTGTGACACCATG 1610
 RESULT 10
 LOCUS AX662357 8473 bp DNA linear PAT 22-MAR-2003
 DEFINITION Sequence 39 from Patent WO02062999.
 ACCESSION AX662357
 VERSION AX662357.1 GI:29163218
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 AUTHORS Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
 Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderina, S.K.,
 Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
 Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
 Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
 Millet, I. and MacDougall, J.R.
 TITLE Proteins and nucleic acids encoding same
 JOURNAL Patent: WO 02062999-A 39 15-AUG-2002;
 Curagen Corporation (US)
 FEATURES
 source
 1..8473
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: 2,42e-60 Length: 8473
 Pred. No.: 619.00 Matches: 111
 Score: 90.67% Conservative: 25
 Percent Similarity: 74.00% Mismatches: 14
 Best Local Similarity: 79.16% Indels: 0
 Query Match: 6 Gaps: 0
 DB: US-10-029-020-14_COPY_1250_1400 (1-151) x AX662357 (1-8473)
 QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
 Db 3600 TACGTGGCGGATATTCCTCTCGAAATGTAAACAGTGTCTTAGAACATAAGAAATAAA 3659
 QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
 Db 3660 GATTTTAGACATAGCAGCAACCCAGCTCATAGATACCTTGCACCGGATCCAGTCACG 3719
 QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
 Db 3720 GGAGATCTGTACGTTCTCGACACAAACACCGCGAGAATTTATCGCCCAAGTCACTTACG 3779
 QY 61 ValVallysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
 Db 3780 GGGGCAAAAGACTTCACTAAAAATGCAAGTGTCTGCGAGGACAGGGAGCAATGCCCTT 3839

QY 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
 Db 3840 CCCTTTGACGAGGCGAGATGTGGGATGAGGGAAGCGCGTGGAGCCACCACTCATGAGT 3899
 QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetile 120
 Db 3900 CCCAAGCAATGGCAGTTGATAAGATGATTAATCTACTTTTGTGTGATGGAACCATGATT 3959
 QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 Db 3960 AGGAAGTTGACCAAAATGGAATCATATCACTCTCTGGGCTCTTAACGATTGTACTTCA 4019
 QY 141 AlaArgProLeuSerCysAspSerValMet 150
 Db 4020 GCCAGACCTTTAACTTTGTGACACCATG 4049
 RESULT 11
 LOCUS AX662359 8487 bp DNA linear PAT 22-MAR-2003
 DEFINITION Sequence 41 from Patent WO02062999.
 ACCESSION AX662359
 VERSION AX662359.1 GI:29163219
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 AUTHORS Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
 Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderina, S.K.,
 Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
 Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
 Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
 Millet, I. and MacDougall, J.R.
 TITLE Proteins and nucleic acids encoding same
 JOURNAL Patent: WO 02062999-A 41 15-AUG-2002;
 Curagen Corporation (US)
 FEATURES
 source
 1..8487
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: 2,42e-60 Length: 8487
 Pred. No.: 619.00 Matches: 111
 Score: 90.67% Conservative: 25
 Percent Similarity: 74.00% Mismatches: 14
 Best Local Similarity: 79.16% Indels: 0
 Query Match: 6 Gaps: 0
 DB: US-10-029-020-14_COPY_1250_1400 (1-151) x AX662359 (1-8487)
 QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
 Db 3596 TACGTGGCGGATATTCCTCTCGAAATGTAAACAGTGTCTTAGAACATAAGAAATAAA 3655
 QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
 Db 3656 GATTTTAGACATAGCAGCAACCCAGCTCATAGATACCTTGCACCGGATCCAGTCACG 3715
 QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
 Db 3716 GGAGATCTGTACGTTCTCGACACAAACACCGCGAGAATTTATCGCCCAAGTCACTTACG 3775
 QY 61 ValVallysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
 Db 3776 GGGGCAAAAGACTTCACTAAAAATGCAAGTGTCTGCGAGGACAGGGAGCAATGCCCTT 3835
 QY 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
 Db 3835 GGGGCAAAAGACTTCACTAAAAATGCAAGTGTCTGCGAGGACAGGGAGCAATGCCCTT 3899

```

Db      3836  CCGTTTGACGAGCGGAGATGTGGGATGGAGGAGCGCGTGAAGCCACACTCATGAGT 3895
QY      101  ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3896  CCCAAGGATGGCAGTTGATGAAGTAAATCTACTTTGTGATGGAACCATGATT 3955
QY      121  ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3956  AGGAAAGTTGACCAAAATGGAATCATATCACTCTTCGGGCTCTAACGATTGACTTCA 4015
QY      141  AlaArgProLeuSerCysAspSerValMet 150
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      4016  GCCAGACCTTTAACTGTGACACCAGCATG 4045

RESULT 12
AX662355
LOCUS      AX662355               8645 bp      DNA      linear      PAT 22-MAR-2003
DEFINITION Sequence 37 from Patent WO02062999.
ACCESSION  AX662355
VERSION     AX662355.1  GI:29163217
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
            Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,
            Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
            Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
            Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
            Millet, I. and Macdougall, J.R.
            Proteins and nucleic acids encoding same
            Patent: WO 02062999-A 37 15-AUG-2002;
            Curagen Corporation (US)
FEATURES    source
            Location/Qualifiers
            1..8645
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      2,48e-60      Length:      8645
Score:          619.00      Matches:    111
Percent Similarity: 90.67%      Conservative: 25
Best Local Similarity: 74.00%      Mismatches: 14
Query Match:    79.16%      Indels:    0
DB:             Gaps:      0

US-10-029-020-14_COPY_1250_1400 (1-151) x AX662355 (1-8645)
QY      1  TyrlleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3772  TAGTCGCGGGATATTCCTCTCGAAATGTAAACAAGTGTCTTAGAAGCTTAAGAAATAAA 3831
QY      21  AspPheArgHisSerHisSerProAlaHisLysTyrTyrlleuAlaThrAspProMetSer 40
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3832  GATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTCACAGGATCCAGTCACG 3891
QY      41  GlyAlaValPheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrVal 60
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3892  GGAGATCTGTACGTTTCTGCACAAACACCCGAGAAATTTATCGCCCAAGTCACATTACG 3951
QY      61  ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3952  GGGGCAAAAGACTTGACTTAAATGACAGAGTCGTCGAGGGACAGGGAGCAATGCCTT 4011
QY      81  ProPheAspThrArgCysGlyAspGlyGlyValAlaThrGluAlaThrLeuThrAsn 100
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      4012  CCGTTTGACGCGGAGATGTGGGATGAGGAGGAGCGCGTGGAGCCACACTCATGAGT 4071
QY      101  ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120

```

```

Db      4072  CCCAAGGAATGCGAGTTGATAAGAATGATTAAATCTACTTTGTGATGGAACCATGATT 4131
QY      121  ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      4132  AGGAAAGTTGACCAAAATGGAATCATATCACTCTTCGGGCTCTAACGATTGACTTCA 4191
QY      141  AlaArgProLeuSerCysAspSerValMet 150
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      4192  GCCAGACCTTTAACTGTGACACCAGCATG 4221

RESULT 13
AX662353
LOCUS      AX662353               8675 bp      DNA      linear      PAT 22-MAR-2003
DEFINITION Sequence 35 from Patent WO02062999.
ACCESSION  AX662353
VERSION     AX662353.1  GI:29163216
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
            Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,
            Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
            Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
            Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
            Millet, I. and Macdougall, J.R.
            Proteins and nucleic acids encoding same
            Patent: WO 02062999-A 35 15-AUG-2002;
            Curagen Corporation (US)
FEATURES    source
            Location/Qualifiers
            1..8675
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      2,49e-60      Length:      8675
Score:          619.00      Matches:    111
Percent Similarity: 90.67%      Conservative: 25
Best Local Similarity: 74.00%      Mismatches: 14
Query Match:    79.16%      Indels:    0
DB:             Gaps:      0

US-10-029-020-14_COPY_1250_1400 (1-151) x AX662353 (1-8675)
QY      1  TyrlleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3784  TAGTCGCGGGATATTCCTCTCGAAATGTAAACAAGTGTCTTAGAAGCTTAAGAAATAAA 3843
QY      21  AspPheArgHisSerHisSerProAlaHisLysTyrTyrlleuAlaThrAspProMetSer 40
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3844  GATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTCACAGGATCCAGTCACG 3903
QY      41  GlyAlaValPheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrVal 60
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3904  GGAGATCTGTACGTTTCTGCACAAACACCCGAGAAATTTATCGCCCAAGTCACATTACG 3963
QY      61  ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3964  GGGGCAAAAGACTTGACTTAAATGACAGAGTCGTCGAGGGACAGGGAGCAATGCCTT 4023
QY      81  ProPheAspThrArgCysGlyAspGlyGlyValAlaThrGluAlaThrLeuThrAsn 100
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      4024  CCGTTTGACGCGGAGATGTGGGATGAGGAGGAGCGCGTGGAGCCACACTCATGAGT 4083
QY      101  ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      4084  CCCAAGGAATGCGAGTTGATGAAGTAAATCTACTTTGTGATGGAACCATGATT 4143

```

121 ArgArgIleAspGlnAsnGlyIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 141 AlaArgProLeuSerCysAspSerValMet 150
 4204 GCCAGACCTTAACTTGTGACACGACGATG 4233

RESULT 14
 AK122513
 LOCUS Mus musculus mRNA for mKIAA1455 protein. 5804 bp mRNA linear ROD 15-MAR-2003
 DEFINITION AK122513
 ACCESSION AK122513.1 GI:28972757
 VERSION FLI CDNA.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Aizawa, H., Yuasa, S., Nakajima, D., Nagase, T., Ohara, O. and Koga, H.
 Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries of DNA Res. 10, 35-48 (2003)
 2 (bases 1 to 5804)
 Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.
 Direct Submission
 Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: mousekazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)
 The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing: Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.
 Location/Qualifiers
 1..5804
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="mbg00055"
 /tissue_type="brain"
 /dev_stage="adult"
 /notes="vector:modified pBC SK+"
 1..5804
 /gene="mKIAA1455"
 <1..5488
 /genes="mKIAA1455"
 /notes="CDS is predicted by in silico analysis. Start codon is not identified."
 /evidence="not experimental"
 /product="mKIAA1455 protein"
 /protein_id="BAC65795.1"
 /db_xref="GI:28972757"
 /translation="LIGVNVFLHSEYGYTTRQDMFLVANGASLTIVFSPGF LTQYHTVWPVNYVYMDTLWMKEENDIPSCDLSGFVPSPIIVSPSTFFRSPSE DSPILPETVOLHEETIPGDLKLSLSSRAAGYKSLKMTQAVIPNLMKVLHMV AVVGRLEKFWFAPSNLAYTFIWDKTDVKNQVGLSEAVSVSGYEVESCLDLTWEK RTVALQGYELDANMGWTKDHLVDVQNGILYKNGENQFISQOQPVVYSIMGNR RRSISCPSCQMGADGNKLLAPVALACIDSLVGDPNVYRIFPSPGNVSVLELRNK DFRHSNPNARYLATDPVGDLYVSDTNTNRIYRPSKLTGAKDLTKNAVAVAGTSEQ CLPFDPEARCGGKGAATEALMSPKMAIDNGLYIFVDGMIRKVDONGHISILGNS RPLWCQVPGVEYVGVHATVETTESATAIYASVSGVLYITETDEKINIRQVTTDGE ISLVAGIIPSECDKNDANCDCYOGDGYAKDAKLNAPSSLAASPDGLYIADLGNRI TAVTDSNKLNSNWFVEASPTDQELVIFDINGTHOYTVSLVTGDXLYNFSYNDNDV GNSGLLATSDSGTWITFFDYDSEGRUYNVTFPGVNVINLHGMKALITVDIESSR EDSVITSNLSIDSFTYTMVDQDLNRSYQIGYDGLSLRIFPYAGSLDSDSHYQTEPHVLGTA

121 ArgArgIleAspGlnAsnGlyIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 141 AlaArgProLeuSerCysAspSerValMet 150
 4204 GCCAGACCTTAACTTGTGACACGACGATG 4233

RESULT 15
 AK122513
 LOCUS Mus musculus mRNA for Ten-m3, complete cds. 8964 bp mRNA linear ROD 08-MAY-1999
 DEFINITION AB025412
 ACCESSION AB025412
 VERSION AB025412.1 GI:4760779
 KEYWORDS Ten-m3.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Aizawa, H., Yuasa, S., Nakajima, D., Nagase, T., Ohara, O. and Koga, H.
 Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries of DNA Res. 10, 35-48 (2003)
 2 (bases 1 to 5804)
 Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.
 Direct Submission
 Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: mousekazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)
 The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing: Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.
 Location/Qualifiers
 1..5804
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="mbg00055"
 /tissue_type="brain"
 /dev_stage="adult"
 /notes="vector:modified pBC SK+"
 1..5804
 /gene="mKIAA1455"
 <1..5488
 /genes="mKIAA1455"
 /notes="CDS is predicted by in silico analysis. Start codon is not identified."
 /evidence="not experimental"
 /product="mKIAA1455 protein"
 /protein_id="BAC65795.1"
 /db_xref="GI:28972757"
 /translation="LIGVNVFLHSEYGYTTRQDMFLVANGASLTIVFSPGF LTQYHTVWPVNYVYMDTLWMKEENDIPSCDLSGFVPSPIIVSPSTFFRSPSE DSPILPETVOLHEETIPGDLKLSLSSRAAGYKSLKMTQAVIPNLMKVLHMV AVVGRLEKFWFAPSNLAYTFIWDKTDVKNQVGLSEAVSVSGYEVESCLDLTWEK RTVALQGYELDANMGWTKDHLVDVQNGILYKNGENQFISQOQPVVYSIMGNR RRSISCPSCQMGADGNKLLAPVALACIDSLVGDPNVYRIFPSPGNVSVLELRNK DFRHSNPNARYLATDPVGDLYVSDTNTNRIYRPSKLTGAKDLTKNAVAVAGTSEQ CLPFDPEARCGGKGAATEALMSPKMAIDNGLYIFVDGMIRKVDONGHISILGNS RPLWCQVPGVEYVGVHATVETTESATAIYASVSGVLYITETDEKINIRQVTTDGE ISLVAGIIPSECDKNDANCDCYOGDGYAKDAKLNAPSSLAASPDGLYIADLGNRI TAVTDSNKLNSNWFVEASPTDQELVIFDINGTHOYTVSLVTGDXLYNFSYNDNDV GNSGLLATSDSGTWITFFDYDSEGRUYNVTFPGVNVINLHGMKALITVDIESSR EDSVITSNLSIDSFTYTMVDQDLNRSYQIGYDGLSLRIFPYAGSLDSDSHYQTEPHVLGTA

121 ArgArgIleAspGlnAsnGlyIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 141 AlaArgProLeuSerCysAspSerValMet 150
 4204 GCCAGACCTTAACTTGTGACACGACGATG 4233

RESULT 15
 AK122513
 LOCUS Mus musculus mRNA for Ten-m3, complete cds. 8964 bp mRNA linear ROD 08-MAY-1999
 DEFINITION AB025412
 ACCESSION AB025412
 VERSION AB025412.1 GI:4760779
 KEYWORDS Ten-m3.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Aizawa, H., Yuasa, S., Nakajima, D., Nagase, T., Ohara, O. and Koga, H.
 Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries of DNA Res. 10, 35-48 (2003)
 2 (bases 1 to 5804)
 Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.
 Direct Submission
 Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: mousekazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)
 The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing: Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.
 Location/Qualifiers
 1..5804
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="mbg00055"
 /tissue_type="brain"
 /dev_stage="adult"
 /notes="vector:modified pBC SK+"
 1..5804
 /gene="mKIAA1455"
 <1..5488
 /genes="mKIAA1455"
 /notes="CDS is predicted by in silico analysis. Start codon is not identified."
 /evidence="not experimental"
 /product="mKIAA1455 protein"
 /protein_id="BAC65795.1"
 /db_xref="GI:28972757"
 /translation="LIGVNVFLHSEYGYTTRQDMFLVANGASLTIVFSPGF LTQYHTVWPVNYVYMDTLWMKEENDIPSCDLSGFVPSPIIVSPSTFFRSPSE DSPILPETVOLHEETIPGDLKLSLSSRAAGYKSLKMTQAVIPNLMKVLHMV AVVGRLEKFWFAPSNLAYTFIWDKTDVKNQVGLSEAVSVSGYEVESCLDLTWEK RTVALQGYELDANMGWTKDHLVDVQNGILYKNGENQFISQOQPVVYSIMGNR RRSISCPSCQMGADGNKLLAPVALACIDSLVGDPNVYRIFPSPGNVSVLELRNK DFRHSNPNARYLATDPVGDLYVSDTNTNRIYRPSKLTGAKDLTKNAVAVAGTSEQ CLPFDPEARCGGKGAATEALMSPKMAIDNGLYIFVDGMIRKVDONGHISILGNS RPLWCQVPGVEYVGVHATVETTESATAIYASVSGVLYITETDEKINIRQVTTDGE ISLVAGIIPSECDKNDANCDCYOGDGYAKDAKLNAPSSLAASPDGLYIADLGNRI TAVTDSNKLNSNWFVEASPTDQELVIFDINGTHOYTVSLVTGDXLYNFSYNDNDV GNSGLLATSDSGTWITFFDYDSEGRUYNVTFPGVNVINLHGMKALITVDIESSR EDSVITSNLSIDSFTYTMVDQDLNRSYQIGYDGLSLRIFPYAGSLDSDSHYQTEPHVLGTA

AUTHORS

Oohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T.,
Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
J. Cell Biol. (1999) In press
2. (bases 1 to 8964)
Oohashi,T.

JOURNAL

Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University.
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
(E-mail: oohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128,
Fax: +81-86-222-7768)

FEATURES

source

Location/Qualifiers
1..8964
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1..8964
/gene="ten-m3"
50..8197
/gene="ten-m3"
/codon_start=1
/product="Ten-m3"
/protein_id="BAA77398.1"
/gi_xref="GI:4760780"

gene

CDS

/translation="MDVKRRPYCSLTKSRREKERRYNTSSADNEECRVPTQKSYSS
ETLKAFDHSDYRLNRYKDLVHREADEYTRQGNFTLRQLGVCEGATRRGVAFCAE
MGLPHRYSISAGSDADTENEAVNPMHWRGVKSGRSSLCSRSNSALTLTP
EHNRSDESEQNNPQPTLOPLPSHKHQAHPHSITSLNNSLTHRNOSAP
PAALPAELQTPPSVOLQDSWLVSNVPLESRHFLFKTGTPPLPSTATPGYTMASG
SVSPPTPRPLNTLSAKFKKSKYCSWRCTALCAVGVSVLLAILLSYFIAMHIF
GLMHLQQTENDTFENGKNSDVTPTNTVSLPSGDNGLKGFTHENNTIDSGEIDIGR
RAIQEVPPIFWRSQFIDQFQELKFNISLQKDALIGYGRKGLPPHSTQYDFVELLD
GTLIAEQRNLVESERAGQARSVLSHEAGFTQYLDGSLTHLAFYNDGNKPNQVSN
TVIESVECPNRCNGECVSGTCHCFGLPGLDCSRAACPVLCNSNGQYSGRCLC
FSGWKTECPDPTQCIDPCGGGIGICMGSACNSYKGENCEADCLDPGCSNHC
CIHGECNFGNGNCEILKTCADQC SGHTVLOESGCTCDPNWTPDCNCEITS
VDCSGHVGMGSCRCGEGWTGACNORACHPRCAEHGCTCKGKCECSQWNGEHCTI
AHYDLKIVGEGCLNSNGRCTLQNGMHCVCQPGWRGAGCDVAMETLCTDSKNEG
DLGIDCMDDPCLQSSCONQPCYRGLPDQDIIISQLOTPSQQAASFYDRIJSLFSG
DSTHVLPGSPFNKSLAVIRGOVLTADGTPPLGVNVSFLHSEYGYTTRQGMEDL
VANGASLTLVFPERSPELTQYHTVIMPNNVYVMDLVMKKEENDIPSCDLSGFVRPS
PIIVSSPLSTFFPSSDESDPIIETQVLHEHTTIPGTDLKLVSRAAGYKSVLKIT
MTQAVIPENLMKHLMAVAVGRULQKWFASPPLAVTTFMDKTDAYNKQYGLSEAVV
SVGEYESCLDLTLMBKRTAVLQGYELDASNMGGWTLDKHVLVDQNGILYKNGENQ
FISQPPVSSIMNGRRRSISCPNGQADGNKLLAPVALACIGISLVGDFNVVR
RIPPSGNVTSVLELRNKRFRHSNPAHYVLAIDPVTGDLVSDTNRIRIYRPSLIG
AKDLTKNAEVAAGTGEQCLPFDEARCDDGKAVEALMSPKMAIDKNGLIIYVDGTM
IRKVDONGIISTLGSNDLTSARPLCTDSMHISQVRLEWPTDILNPMDSNIYVLDN
NVVLQITENQVRIAGRMHCQVPEYVGHKAVQTTLESATIAVSYSGVLIITE
DEKKLNRQVTTDEISLVAGIPSECDKNDANDCYQSGDGYAKDAKLNAPSILA
ASPDGTLYADLGNIRIRAVSKNKLPLNMGNFYEVASPTDQELYIFDINGTHQYTVSL
VTGDLVNFYSYNDNDVATVDSNGTLAIRRDENMPVRYVSPDNQVTLILGNGC
LKSMQAQLELVLFTVHGNSGLLAKSDETGTTTFDYSBGRNLNVTTFVGVTNLH
GMDKALTVDIJESSREEDVSTNSLSSIDSTFTMVQDQLRNSYQIGDYSGLRIFYAS
GLDSYQTEPHVLGANTANPVAKRNMTPENGQNLVEWRFRKQAQKVNVTGKLR
VNGRLLSVDFDTTEKTEKIYDDHFKFLRIAYDTSGHPTLWLPSPKLMANVTYSST
GGQIASIQRTGTTSEKVDYDQGRIVSRVADGKTWSYTYLEKSWILLHSHQOYIPEYD
MMDRLSITMPSVARHTMOTIRSIGYRNINYPNESNASIITDYNEEGLLQTAELGT
SRRLVFKRQTRLSEILYDSTRVSFTYDAGVLKTNLQSDGFICTIRYQIGPLI
DRQIFPESDGMVNAFDSYDNSFRVTSMQGVNETPLPIDLYQFDIDISGVEQFGK
FKIGEPANTTKAYEYVDVQLOQTVYLNKIMRYNYDLNGLNHLNPSSSARLTPLR
YDLDRITRLGVQVRLDDEGFLRGRTIEFYSGKGLLTRVYSKSGWTVIYRVDGL
GRVSSKTSLGQLOFFVADLYTPTRI THVYNSSEITSLYDIQGHLPAMEISSGD
EYFIASDNTGTPLAVFSSNGLMKOTQYTAIGEIFYDSNVDRQLVIGHGGLYDEPLTK
LIHFGERDYDILLAGRWTPDIDIEIKRIGKDPAPFLNYPNNNNPASKHVDKDYITD
NSWLVTFGPHLNAIPGFVPVKDTEPSYELVKVSGDEWVPIFEGVQOVARQAKAF
LSLGGKAEVQVSRKAGAEQSWLWFATVKSLKGVMLAVSQGRVQTNVLNLANEDCI

ORIGIN

Alignment Scores:

Pred. No.: 3.37e-60 Length: 8964
Score: 618.00 Matches: 110
Percent Similarity: 90.67% Conservative: 26
Best Local Similarity: 73.33% Mismatches: 14
Query Match: 79.03% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AB025412 (1-8964)

QY 1 TyrIleArgArgIlePheProSerClyAsnValThrAsnIleLeuGluLeuArgAsnIlyS 20
Db 3653 TACGTCCGGCGGATATCCCGTCTGGGAATGTGACAGTGTTTAGAACTAAGAAATAA 3712
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3713 GATTTAGACATAGTAGCAACCCAGCTCACAGATACCTACCTGGCTACGACCCAGTCACC 3772
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLeuSerThrVal 60
Db 3773 GGAGATTGTACGTCTCTGATCTAATACACCGCGAGAATCTATCGCGCGAATCATCTACG 3832
QY 61 ValValLysAspLeuValLysAsnSerGluValValalaglyThrGlyAspGlnCysLeu 80
Db 3833 GGAGCCAAAGACCTGACTAAAAACGCTGAAGTGTGTGGCAGGACCGGGGAACAGTGCCTT 3892
QY 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 3893 CCCTTTGACGAGCCAGGTGTGGGATGGAGGCAAGCTGTGGAGCAACCGCTCATGAGT 3952
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 3953 CCCAAAGGAATGCAATCGATAAGAACGGACTGATCTACTTTGTGTGGAACCACTGATC 4012
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4013 AGAAGTTGTCATAAATGGAATCATATCAACTCTCTCTGGGCTCCACAGCACCTCACGTCA 4072
QY 141 AlaArgProLeuSerCysAspSerValMet 150
Db 4073 GCTCGACCTTTAACCTGTGATAGCATG 4102

Search completed: August 14, 2004, 11:52:42

Job time : 2702.71 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 4616.26 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_1490_1750

Perfect score: 1384

Sequence: 1 KINRIQVTTSGEISLVAGA.....SSKDDVTITNLSAGAFYT 261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn.model -DEV=xlh
-Q=/cgm2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgm2_1_19065@runat_06082004_112216_29275 -NCPV=6 -ICPV=3
-NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.roi.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | | | DB | ID | Description |
|------------|--------|-------|--------|----|-----------|-------------|
| | Score | Match | Length | | | |
| 1 | 1384 | 100.0 | 8354 | 6 | AX556500 | Sequence |
| 2 | 1384 | 100.0 | 8438 | 6 | AX675551 | Sequence |
| 3 | 1384 | 100.0 | 8993 | 9 | HSN808812 | Sequence |
| 4 | 1380 | 99.7 | 8645 | 6 | AX600210 | Sequence |
| 5 | 1370 | 99.0 | 8585 | 10 | AB025413 | Mus muscu |
| 6 | 1361 | 98.3 | 9722 | 10 | AF059485 | Mus muscu |
| 7 | 1203.5 | 87.0 | 8624 | 9 | AB037723 | Homo sapi |
| 8 | 1130 | 81.6 | 3048 | 9 | AK127705 | Homo sapi |
| 9 | 1112 | 80.3 | 112872 | 2 | AF000716 | Homo sapi |
| 10 | 1112 | 80.3 | 163812 | 2 | AF001141 | Homo sapi |
| 11 | 1112 | 80.3 | 170156 | 2 | AF002412 | Homo sapi |
| 12 | 1112 | 80.3 | 186084 | 9 | AP002768 | Homo sapi |
| 13 | 1103 | 79.7 | 9264 | 5 | AB026980 | Danio rer |
| 14 | 1098 | 79.3 | 226256 | 10 | AC073599 | Mus muscu |
| 15 | 1091 | 78.8 | 245722 | 2 | AC120288 | Rattus no |
| 16 | 1091 | 78.8 | 257693 | 2 | AC103323 | Rattus no |
| 17 | 943.5 | 68.2 | 5804 | 10 | AK122513 | Mus muscu |
| 18 | 943.5 | 68.2 | 7816 | 10 | AF195418 | Mus muscu |
| 19 | 943.5 | 68.2 | 8964 | 10 | AB025412 | Mus muscu |
| 20 | 939.5 | 67.9 | 5309 | 9 | AB040888 | Homo sapi |
| 21 | 935.5 | 67.6 | 8473 | 6 | AX662357 | Sequence |
| 22 | 935.5 | 67.6 | 8487 | 6 | AX662359 | Sequence |
| 23 | 935.5 | 67.6 | 8645 | 6 | AX662355 | Sequence |
| 24 | 935.5 | 67.6 | 8675 | 6 | AX662353 | Sequence |
| 25 | 935 | 67.6 | 5583 | 10 | AK122490 | Mus muscu |
| 26 | 917 | 66.3 | 8118 | 5 | GA238613 | Gallus ga |
| 27 | 914 | 66.0 | 8373 | 10 | AB025410 | Mus muscu |
| 28 | 898 | 64.9 | 8297 | 9 | AF100772 | Homo sapi |
| 29 | 889 | 64.2 | 193544 | 5 | AL807754 | Zebrafish |
| 30 | 884.5 | 63.9 | 8816 | 5 | AB026979 | Danio rer |
| 31 | 865.5 | 62.5 | 8689 | 6 | AX250067 | Sequence |
| 32 | 865.5 | 62.5 | 8689 | 10 | AF086607 | Rattus no |
| 33 | 864.5 | 62.5 | 8797 | 6 | AX250063 | Sequence |
| 34 | 864.5 | 62.5 | 8797 | 6 | AX250066 | Sequence |
| 35 | 864.5 | 62.5 | 8797 | 10 | AB025411 | Mus muscu |
| 36 | 863.5 | 62.4 | 6560 | 6 | AX250065 | Sequence |
| 37 | 863.5 | 62.4 | 7713 | 9 | HSN808325 | Homo sapi |
| 38 | 863.5 | 62.4 | 7781 | 9 | AB032953 | Homo sapi |
| 39 | 863.5 | 62.4 | 9729 | 6 | AX250013 | Sequence |
| 40 | 863.5 | 62.4 | 9826 | 6 | AX250008 | Sequence |
| 41 | 862.5 | 62.3 | 8409 | 5 | GA279031 | Gallus ga |
| 42 | 862.5 | 62.3 | 8409 | 6 | AX250068 | Sequence |
| 43 | 861.5 | 62.2 | 8575 | 6 | AX921803 | Sequence |
| 44 | 794 | 57.4 | 229533 | 2 | AC106382 | Rattus no |
| 45 | 794 | 57.4 | 321049 | 2 | AC137221 | Rattus no |

ALIGNMENTS

RESULT 1

AX556500
LOCUS AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gangolli, E.A., Patturajan, M., Vernet, C.A., Malyanankar, U.M.,
Kekuda, R., Stone, D.J., Anderson, D., Shimkets, R.A., Burgess, C.E.,
Zernhusen, B.D., Liu, X., Spytek, K.A., Casman, S.J., Boldog, F.L.,
Smithson, G., Li, L. and Ji, W.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
FEATURES
source 1. 8354
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.92e-128 Length: 8354
Score: 1384.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x AX556500 (1-8354)
Qy 1 LysIleAsnArgIleAArgGlnValThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4502 AAGATCAACCGCATCAGCGAGGTCAACCATAGTGAGAGAGATCTCACTCGTTGTGGGGCC 4561
Qy 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4562 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATGTTTTCTGGAGACCATGGT 4621
Qy 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4622 TATGCCAAGGATGCAAGTTAAATACCCCATCTTCCTGGCTGTGTCTGATGGGGAG 4681
Qy 61 LeuTyrValAlaAspLeuGlyAsnIleArgPheIleArgLysAsnLysProPhe 80
Db 4682 CTCACGTGGCCGACCTGTACACCAAGCCCTGCCACAGAGACTACCTCTGTACAC 4741
Qy 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4742 CTCAACACCCAGAACATGATAGCTGTCTTCAACAAATGACCAAGGAGCTCTATCTGTT 4801
Qy 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4802 GATACCAACCGCAAGCACTGTACACCAAGCCCTGCCACAGAGACTACCTCTGTACAC 4861
Qy 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4862 TTCCTTACACTGGGGACGGCAATACACTCATCAGACACAAATGGCAACATGGTA 4921
Qy 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4922 AATGTCCGCCAGAGACTCTACTGGGATGCCCTCTCTGGTGGTGGTCCAGAGTGGCCAGGT 4981
Qy 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4982 TACTGGGTGACANTGGGCACCAACAGTGCATCAAGAGTGTGACCAACACAGGCACAG 5041
Qy 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200


```

QY      81  LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db      4748  CTCACACCCAGAACATGATGAGTGTCTTACCAATGACACAGAGCTCTATCTGTTT 4807

QY      101  AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db      4808  GATACACCGCAAGCACCTGTATACCCAAAGCTGCCACACAGGACTACCTGTACAC 4867

QY      121  PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db      4868  TTCACCTACACTGGGACCGGACATCACATCATCACACAAATGCAACATGGTGA 4927

QY      141  AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
Db      4928  AATGTCGCGGAGACTCTACTGGAGTCCCTCTGGCTGGTGTCCAGATGCCAGGTG 4987

QY      161  TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db      4988  TACTGGGTGACCATGGGACCAACAGTGCACCTCAAGAGTGTGACCACACAGGACAG 5047

QY      181  LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db      5048  TTGGCCATGATGACATACCATGCAATCCGGCTTCTGGCAACCAAAAGCAATGAAAC 5107

QY      201  GlyTyrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db      5108  GGATGGACCAACATTTATGATGACGACACTTGGCCGCTTGACAAATGTGACCTTCC 5167

QY      221  ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db      5168  ACTGGCCAGTGAGGAGTTTCCGAGTGATACAGACAGTTCAGTGCATGCCAGGTAGAG 5227

QY      241  ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db      5228  ACCTCCAGCAAGGATGATGATCACCATAACCAACCTCTGCTCCTCAGCGCTTCTAC 5287

QY      261  Thr 261
Db      5288  ACA 5290

RESULT 3
HSM806812      8993 bp      mRNA      linear      PRI 28-AUG-2003
LOCUS          Homo sapiens mRNA; cDNA DKFZp686K11107 (from clone DKFZp686K11107).
DEFINITION    BX640737
ACCESSION     BX640737.1 GI:34364828
VERSION       Homo sapiens (human)
KEYWORDS      Homo sapiens
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 8993)
AUTHORS      Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Wewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
CONSTRM      The German Human cDNA Consortium
TITLE         Direct Submission
JOURNAL       Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberger, GERMANY
COMMENT       Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686K11107) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
LOCATION/Qualifiers
FEATURES      1..8993
source        /organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/map="111q14.1"
/clone="DKFZp686K11107"
/tissue_type="human fetal kidney"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="fetal"
1..8993
/gene="DKFZp686K11107"
<1..4200
/gene="DKFZp686K11107"
/Note="hypothetical protein, N-terminus elongated,
differentially spliced"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAE45850.1"
/db_xref="GI:34364829"
/translation="RIDQNGIILSTLLGNDLTSARPLSCDSVMDISQVHLEWPTDLAI
NPMNSLVLDNNVLIQISENHQVRIVAGRMHCQVPGIDHDFLLSKVAIHATLESAlA
LAVSHGLVLYIAETDEKIKINIRQVTTSGEISLVAGASGCDKNDANDCDFSGDDGY
AKDAKINTPSSLAVCAGDELXYADIGNIRIRKKNKPFINTQNMVELSSPIDQELYL
FDTTGKLYTQSLPTGDLXNYFTYDGDITLITDNGNMVNRDSTGMLWLVVPD
GVYVWTGNTSALKSVTTQGHELAMWTYHNSGLLATKSNENGWTTFYEDSRRLT
NVTFTGVSSPSPDTSVHVQVETSSKDDVTITNLSASGAFYTLIQDQVRNYYII
GADGSLRILLANGMEVALQTEPHLLAGVNPVGRKNVTLPIDNGLNLVEMQRKEQA
RGQVTFGRURKVRHNRLSLUDFDRVTEKIIDHDKFTLRILKYDQAGRPLSKWPS
RINGVNTVSPGGYTAGIQRGIMSERMEYDQAGRITSRIFADGKTSVYTLSEKMWLL
LHSQRYIFEPDKNDRLSVTPNPVARQTLTIRSVYRNIYQPEGNASVIOFTE
DHLLHTFVLGTGRVIVYKGLSKLAETLYDTTKVSYTDTAGMLKTLINQSGFT
CTIRVOIGPLIDROI FRFTREGMVNABPDVYDMSFRVTSQAVINETPLPIDLYRY
DDVSKTQFGKGVYVDINQIITAVMTHKFDAYGRMKEVQTEIFRSLMYWMTV
QIDNMRVKKELGVYDNPITRYSYDADGQVTSINDKPLMRYSDRSLNGNLEHL
SPGNSARILTPLRIDIRITRLGDVQYKMEDEFLRQGGDIFEYNSAGLLIKYNRA
GSWSVRYVDGLGRVSSKSHHLOFFADLTNPVKVHLNHSSTITSLYDLOQ
GHLFAMELSSGDEFYIACDNIGTPLAVSGTGLMIKQILYTAGVGLYMDTDFNFOII
GVHGLYDPLTKLVEMGRDYDLAGRWTSPDHELWKHLSSNVWPFNLYMFKNNPI
SNSQIKFMTDVSNLWTLFGQLHNVIIPGYPMDAMEPSLELITQMTQOEMDNS
KSILGVQCEVQKQKAFVTLERFDLGYSTITSCQAPKTKFASGSGVFGKVPAL
KDGRTVDIIISVANDGRVAAILNHAHYLENLHTIDGVDTHFYFKPGPSGLAIL
GLSGRRITLNGVNVTVQINTVLNGRTRRYTDILOQYCALCLNTRYGTTLDEEKARV
LELARAVRQAWAREQQRLREGERGLRWTGEGKQVLSLSTGRVQYDGFVSVSEVQ
PELSDSANIHFMRQSEMR"
8953..8958
/gene="DKFZp686K11107"
8973
/gene="DKFZp686K11107"

polyA_signal
polyA_site

ORIGIN
Alignment Scores:
Pred. No.:      3,21e-128      Length:      8993
Score:          1384.00      Matches:      261
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              9          Gaps:      0

US-10-029-020-14_COPY_1490_1750 (1-261) x HSM806812 (1-8993)
QY      1  LysileAsnArgileArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db      358  AAGATCAACCCGATCAGCAGCGTACCACCTAGTGAGAGATCTCACTCGTGTCTGGGCC 417

QY      21  ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db      418  CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTTCTTCTGGAGACCATGT 477

QY      41  TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db      478  TATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTGTGTGATGGGAG 537

QY      61  LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80

```

Db 538 CTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTC 597
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProfilAspGlnGluLeuTyrLeuPhe 100
Db 598 CTCACACCCAGAACATGATAGCTGTCTTCACCAATTCACGAGGAGCTCTATCTGTTT 657
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 658 GATACACCCGCGAAGACCTGTACACCAAGGCTGCCACAGAGACTTACTGTACAAC 717
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuLeuThrAspAsnAsnGlyAsnMetVal 140
Db 718 TTCACCTACACTGGGAGCGGCGACATCACACTCATCACAGACCAACAAATGGCAACATGGTA 777
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTripleuValValProAspGlyGlnVal 160
Db 778 AATGTCCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGTG 837
QY 161 TyrTriPValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 838 TACTGGGTGACCATGGCCACCAACAGTGCCTCAAGAGTGTGACCAACAGGACACGAG 897
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 898 TTGGCCATGATGATACATACCATGGCAATTCGCGCCTTCTGGCAACCAAAAGCAATGAAAC 957
QY 201 GlyTriPThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 958 GGTATGGACACATTTTATAGTAGTACGACAGCTTTGGCGCCTTGACAAATGTGACCTTCCT 1017
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 1018 ACTGGCCAGGTGACAGCTTTCGAGTGATACAGACAGTTCAGTGCATGCCAGGTAG 1077
QY 241 ThrSerSerLysAspAspValThrLleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 1078 ACCTCCAGCAAGATGATGTACCATAAACCAACCACTGTCTGCTCAGGCGCCTTCTAC 1137
QY 261 Thr 261
Db 1138 ACA 1140

RESULT 4
AX600210 8645 bp DNA linear PAT 14-FEB-2003
LOCUS
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S.,
Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,
Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K.,
Khare,R. and Walia,N.K.
TITLE Proteins associated with cell growth, differentiation, and death
JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1 .8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7488573CB1"
ORIGIN
Alignment Scores:
Pred. No.: 7,69e-128 Length: 8645
Score: 1380.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0

Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.71% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x AX600210 (1-8645)
QY 1 LysileAsnArgileArgGlnValThrThrSerGlyGluLeuSerLeuValAlaGlyAla 20
Db 4551 AAGATCAACCGCATCAGGACAGGTCCACACTAGTGGAGAGATCTCACTCGTTGCTGGGGCC 4610
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4611 CCCAGTGGCTGTGACTGTAAATAATGATGCCAATGTGAATTTCTTCTGGAGACGATGGT 4670
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4671 TATGCCAAGATGCAAGATTAAATACCCCATCTTCTTGGCTGTGTGTGTGTGGGAG 4730
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgileArgPheileArgLysAsnLysProphe 80
Db 4731 CTCTAGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTC 4790
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProfilAspGlnGluLeuTyrLeuPhe 100
Db 4791 CTCACACCCAGAACATGATGAGCTGTCTTCCAAATTCACCAAGGAGCTCTATCTGTTT 4850
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4851 GATACACCCGCGAAGACCTGTACACCAAGGCTGCCACAGAGACTTACTGTACAAC 4910
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuLeuThrAspAsnAsnGlyAsnMetVal 140
Db 4911 TTCACCTACACTGGGAGCGGCGACATCACACTCATCACAGACCAACAAATGGCAACATGGTA 4970
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTripleuValValProAspGlyGlnVal 160
Db 4971 AATGTCCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGTG 5030
QY 161 TyrTriPValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5031 TACTGGGTGACCATGGCCACCAACAGTGCCTCAAGAGTGTGACCAACAGGACACGAG 5090
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5091 TTGGCCATGATGATACATACCATGGCAATTCGCGCCTTCTGGCAACCAAAAGCAATGAAAC 5150
QY 201 GlyTriPThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5151 GGATGGACACATTTTATGAGTAGCAGAGCTTTTGGCCGCTTGACAAATGTGACCTTCCT 5210
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 5211 ACTGGCCAGGTGACAGCTTTCGAGTGATACAGACAGTTCAGTGCATGCCAGGTAGAG 5270
QY 241 ThrSerSerLysAspAspValThrLleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 5271 ACCTCCAGCAAGATGATGTACCATAAACCAACCACTGTCTGCTCAGGCGCCTTCTAC 5330
QY 261 Thr 261
Db 5331 ACA 5333

RESULT 5
AB025413 8585 bp mRNA linear ROD 08-MAY-1999
LOCUS
DEFINITION Mus musculus mRNA for Ten-m4, complete cds.
ACCESSION AB025413
VERSION AB025413.1 GI:4760781
KEYWORDS Ten-m4.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1 (sites)
Ohashi, T., Zhou, X., Feng, K., Richter, B., Moergelin, M., Perez, M.T., Su, W., Chiquet-Ehrismann, R., Rauch, U., and Fassler, R.
Mouse Ten-m/Odz is a new family of dimeric type II transmembrane proteins expressed in many tissues
J. Cell Biol. (1999) In press
2 (bases 1 to 8585)
Ohashi, T.
Direct Submission
Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho, Okayama, Okayama 700-8522, Japan
(E-mail: ohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128, Fax: +81-86-222-7768)
Location/Qualifiers
1..8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1..8585
/gene="ten-m4"
188..8503
/gene="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"
/translation="MDVKEKPYRSLTRRDAERYTSSADSEKGGKPKQSYSSST LKAYDQDARLAYGKVDKMPQEAEEFCRTGTNFTLRELGLCEMTPHGTLRTDGL PHCYGMSGASDADLEADTLSPHPVRLWGRSTRSGSSCLSSRANSNLILDTTEHE NTEHDHPSLQNHPLRPPPLPHAHPTQHHAAINSLNKGNFTSPNSPAPTDH SLUGPPAGSAGETHAQNMLLSNIPLTEHLKQPLPLGLTDNLNLEMDILSARH DGYSDHFLFKFPGSPFLCTSPCTSPYTSVSPPLRPLRSPRFPAPNLKAPS KYCNWKAALSAIATLIVLAVFAMHLFGLNHLQPMGOMQWETIDTASW PVPDVSILPDSGTGLETPDRKGAAKGRPSLPEDSFINDSEIDVGRASOKIP GTFMSOVFIHPVHLKENVSLGKAAVLGIRKGLPSHTQFVELLDGRLLTQE ARSLGQRQSRGVPFPPSHETGFIQYLDGSLWHLAFYNDGKESVFLTVAISVD APCNYGNGDCISGTCHFLGFLGDCGRASCPVLCSGQYMKGRCLCHSGWKGA ECDPTNQCIDVACSHGTCIMGTGICNFGYKGESCEVDMDPTCSSRVCVGRGCH SVWGTCNCTPRATCLDQCSGHGTFLPDTGLCNDPSPWTHDCSIEICAAACGHHV CVGTCRCEDWMAACDQRAHPRCAEHGTCDGCKCSPGWNGEHCITLHYLDREV KEGCPGLNGRCCTLDLNGHVCOLGWRGTGCTSMETCGDKDNDGGLVDCMD PDCCQLCHVNPILCSFDPLDIIQETQAPVSQNLNAPFFDIRKFLVGRDSTHSIP ENPDGHACVPLGQVMTSDGTLVGNVINSFINNPLFGYTSRQDGSFIDLVTNGGSI ILRPERAPFITOEHLWLPDRFFVMTIYMRHEENIPSCDLNFPARNPVPSPSL TSFASCAEKGEIPEIQALQBEIYIAGCKMRLSYLSRTPGYKSVLRLSLTHPTIP NLMKVHLVAVGRLFRKFAAAPLSYVFTWDTVDVYNOKVFGSPFAFVSVEYES CPDLILWEKRTAVLQGYEIDASKLGGWSLDKHHALNIOGSLHKGNGENQVSOQPV IGSIMNGRRSISCPSCNGLADGNKLLAPVALTCGDSGLYVGFDFNIRIFPFGV TNLNMRNKPFRHSPHAKXYLATDPMGAVFLSDTNSRRVFKVSTTVVKDLVNS EVAGTGDCQPLDTRCGDGKATEATLTPRGITVDKFLIYFVDGTMRIRVDONG IISLLGSDNTSARPLSCDSVMEISQVRLWEPDTLAINPMDNSLYVLNDNNVLIQISE NHQVRI VAGRMHPCVOPGIDHFLSKVAIHALESATALAVSHNGVLVIAETDEKIN RIQVTTSGEISLVAGAPSGCDKNDANCDFSGDDGYAKDKLNTPSLAVCADGEL YVADLGNRIPIRKKNPPLNTQWYELSPIDQELYLFDISGRHLYTQSLPDGILY NFTYTGDDIITHIDNNGVNRDSTGMPLVVPDQVYVMTGNSALRSVTQ GHELAMTYHGNSGLLTAKSNGWTFTEYDFSRGLNTNFTPTGOVSFSDSDTSV HVQVETSCKDVTITNLASAGAFYTLQDQVRSYIYGADGSLRLILANGMEVALOT BPHLLAGTVNTPVGRNVTPLDINGLNLVWQRKEQAGOVTVFGRRLVHNRNLS LDQFVTRTEKIYDHRKFTLIDYQAGRPSLSPSSRLNGVNVTVSPGGIAGIQR GIMSERMEYDQAGRTSIFADGKMSYTYLEKSNVHLHSQRYIIEFDKNDRLSV TNPVNRQTLETISVGYRNIYQPEGNASVIOFTEDHLLHTFLYLGTRVYKY GKLKSLAETLDTKVSFYDTEAGMLKTNLQNEGFTCTTIRYQILGLDIRIFRT EGMVNAFNDYNSFRVTSMQAVINETPLIDLYRYDDVSGTEQFGKFLVYDI NQIITATWHTHFDAYGRMKVQYEIFRSLMYMTVOYDNGMGRVVKMKGVYAN TTRYSEYVDAQOLQTSINDKPLWRYSDYDNGNLHLLSPGNSARLLPLRLDLRORIT RLGDVOYKMDGELQRQGDVFNENAGLLIKAYNRASGWSVRTRYDGLGRRVSSKS SHSHLQFFYADLNTPTKTVLHNSSEITSLYDYLQGLFAMELSSGDEFFIACDN IGTPLAVFSGGLMKILQIITYAYGEIYMDTNPFIIGYHGGYLDPLTLVHMGRRD YDVLAGRWTSPDHLWLKELSSNIVPFLHYMFKNNNPISQDIKCFMTDVSMLLT

ORIGIN
Alignment Scores:
Pred. No.: 7,68e-127 Length: 8585
Score: 1370.00 Matches: 258
Percent Similarity: 99.62% Conservative: 2
Best Local Similarity: 98.85% Mismatches: 1
Query Match: 98.99% Indels: 0
DB: 10 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x AB025413 (1-8585)
QY 1 LysilleAsnArgilleArgGlnValThrThrSerGlyGluileSerLeuValAlaGlyAla 20
Db 4661 AAGATCAATCGCATCAGGCGAGTCACTACAAGTGTGTGAGATCTCACTGCTGTGCTGCTG 4720
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4721 CCCAGTGGCTGTGACTGTAAATAATGATGCCAATGTGACTGTCTTCTCTGGAGATGATGT 4780
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4781 TACGCCAAGGATGCAAGAGTGAATACCCCATCGTCTCTTGGCTGTGTGTCTGACGGGAG 4840
QY 61 LeuTyrValAlaAspLeuGlyAsnLeuArgilleArgPheilleArgLysAsnLysProPhe 80
Db 4841 CTCCTATGTGGCCGACTCGGAAACATCCGAATTCGATTTATCCGGAAGAACCAAGCCTTTC 4900
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4901 CTGAACACTCAGAACATATGATCAGAGGTATCTCTCCCATCGACAGAGCTGTACCTCTTT 4960
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4961 GATACAGTGGCAAGCATCTGTATCACTCAGAGCTTACCCACAGGGGACTTACCTGTATCAAC 5020
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuThrAspAsnGlyAsnMetVal 140
Db 5021 TTCCTTACACAGGGGACGGGACATCACATATCACCGACACCAATGGCAACATGCTGTG 5080
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 5081 AACGTCCGCGGAGACTTACCGGGATGCTCTCTGGCTGTGTAGTCCACAGATGGCCAGGTA 5140
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5141 TACTGGGTAAACATGGGACCAACAGCGCACTCAGAGAGTGTGACCAACAGGACACAGG 5200
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5201 CTAGCCATGATGACCTACCATGGCAACTCTGGCTCTTGGCAACCAACCAAGCAATGAAAC 5260
QY 201 GlyTrpThrThrPheTyrGlyTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5261 GGGTGGACAAAGTTTATGATGATGACAGTTTGGTCCCTGTGACCAACAGTCACTTTCCA 5320
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 5321 ACTGGCCAGGTGAGCAGTTTCCGAAGCGATACACAGCAGCTCAGTGCAGCTGAGTAGAG 5380
QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 5381 ACCTCAGCAAGATGACGTACCATTAACCAACCAACCTGTCTGTCTGGGTGCTCTTCTAC 5440
QY 261 Thr 261
Db 5441 ACC 5443

```
RESULT 6
AF059485      9722 bp      mRNA      linear      ROD 15-AUG-1998
LOCUS      Mus musculus DOC4 (Doc4) mRNA, complete cds.
DEFINITION      Mus musculus DOC4 (Doc4) mRNA, complete cds.
ACCESSION      AF059485
VERSION      AF059485.1 GI:3170614
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1. (bases 1 to 9722)
Wang,X.-Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P.,
Zinsner,H. and Ron,D.
Identification of novel stress-induced genes downstream of chop
EMBO J. 17 (13), 3619-3630 (1998)
MEDLINE      98315054
PUBMED      9649432
REFERENCE
2. (bases 1 to 9722)
Wang,X.-Z. and Ron,D.
Direct Submission
Submitted (14-APR-1998) Skirball Institute, New York University
Medical Center, 550 First Ave., New York, NY 10016, USA
FEATURES
source
1..9722
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/cell_line="NIH-3T3"
/notes="tunicamycin-treated"
1..9722
/gene="Doc4"
/notes="downstream of CHOP 4"
583..9060
/gene="Doc4"
/notes="similar to Drosophila melanogaster tenm/odz and
human gamma-heretulin; type II transmembrane protein"
/codon_start=1
/product="DOC4"
/protein_id="AAC31807.1"
/db_xref="GI:3170615"
/translators="MDVKKRKYSLRRDARRYTSSSADSEGGKPKSYSSST
LKAYDQARLAYGSRVMDVMPQAEERCRGTNFTLRELGLGEMTPRHGTLRTDGL
PHCSYMGASDADLEADTVLSPHPVRLWGRSTRSGRSCLSRANSNLTLDTEHE
NTEGAPLHCSASSTPIEQSPSPPPANESQRRLLGNVAQDLPDSSESEFVEN
SFLVKGASLGAANDHPSLSQNHPLRTPPPPLPHATPNQHHASINSLNKGNFT
PRNPSAPTDHSLSGPPAGSAQEPHADQNVLSKIPIVETRNLGKQFPLGTWQDN
LIEMDIFASRRDGYSDGHFFRPGGTPLFCTSPCYPLTSTVSPPPPLPRST
FSRPAFLKPKSKYCNKCAALSAILISATLIVLLAVFAMHLFGLNHLQPMGQM
MYELTEDTASWPTDVSILYPSGGTGLETPDRKGKGAEGKPSLPEPSFIDSGEI
DVGRASQKIPGTFWRSQVFIIDHPVHLKFNVLGKAAVGIYGRKGGLPPSHQDFFV
ELIDGRLRLQEARSLRGPQRQSPVPPSSHETGFTQYLDSDGHLWLAFYNDGKESEV
VSFLTATIESDNCPCNYGNGDCISGTCFGLGFLGDCGRASCPLVCSGNGKMG
RCLCHSKGAECQVPTNOCIDVACSHGTCIMGTICNPGYKGSCEBVDMDPTCS
SRGVCPGCHCVSGWGTNCTPRATCLDQCSHGTLPLDTGLCNDPMSWTHDCSI
EICADQCGHGVCGTCREDDWMAACDQACHPACAEHGTCDGKCECTPCWNE
HCTIEGCPGNGRGTLDLNGHVCVQLGWRGTGDTSMETCGCGKNDGDLVD
CMDDPCLQPLCHVNPCLGSPDLIIQETQAPVSQNLNSFYDRKFLVGRDSTHS
IPGNPDGCHACVIRGQVMTSDGPLVGVNIFSPNNPLFGYTI SRQDSFDLVNNG
ISILIRERAPFIQETLWLPDRFFVMEITVMRHEENIPSRDLNFAFPNVPSP
SPLTSPASSCAEKGPVPEIQALQBEIVIACTMKRLSYLSRTPGYKSVVRLISLTHPT
IPNIMKVLWAVEGFLFRKFAAADLSVYFTWDTVDVNOVFLSFAFVGVGE
YESPCDLINKERRAVLQGYEDASKLGGSLDKHHLNIOGSLHKGNGENQFVSQO
PPVIGLNGNRRTSLCPCSNADGNKLAPALVATCGSGSLYVGFNVRIRFPSP
GNVNILEMSPHAKYLLATDPMGSAVFLSDTNSRRVKFVITTVKDLVKNSEVVA
GTGQCQLPFDTRCDGGKATEATNPRGTVDKFGLIYFVDGTMRRVQNGIIST
LGGNDLTPARLSCDSVMEISQVRLEWPTDLAINMDNSLYLDNINVLQISNNHOV
RIVAGRPMCOVRPGDIFLLSKVAIHTLESATLAVSHGVLYIAETDEKKINRQO
VITSGEISLVAGSGCCKNDANCDFSGDGYAKAKLNPSSLVACADGELYVAD
LGNIRIRIRKRNPLNTQNMFLSPDQELYLPTDSGKHLVQSLPGLVLYNTY
TGDGDITHITONNGMNVNRDSTGMPLWLVPVPGQVYVMTNGINSALRSYLTQGH
```


Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 111232 bases at least Q40
 Consensus quality: 111769 bases at least Q30
 Consensus quality: 112045 bases at least Q20
 Insert size: 112172; sum-of-contigs
 Quality coverage: 13.62x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 36854 contig of 36854 bp in length
 36955 60659 contig of 23705 bp in length
 60760 75582 contig of 14823 bp in length
 75683 89456 contig of 13774 bp in length
 89557 99398 contig of 9842 bp in length
 99499 105224 contig of 5726 bp in length
 105325 108843 contig of 3519 bp in length
 108944 112872 contig of 3929 bp in length.

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 36854: contig of 36854 bp in length
 36955: gap of 100 bp
 60659: contig of 23705 bp in length
 60760 75582: gap of 100 bp
 75683 89456: contig of 14823 bp in length
 89557 99398: gap of 100 bp
 99499 105224: contig of 13774 bp in length
 105325 108843: gap of 100 bp
 108944 112872: contig of 9842 bp in length
 108944 112872: contig of 5726 bp in length
 108944 112872: contig of 3519 bp in length
 108944 112872: contig of 3929 bp in length.

FEATURES

source
 1. .112872
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q14"
 /clone="CMB9-72M1"
 1. .36854
 /note="assembly_fragment"
 36955 .60659
 /note="assembly_fragment"
 60760 .75582
 /note="assembly_fragment clone_end:SP6 vector_side:left"
 75683 .89456
 /note="assembly_fragment clone_end:T7 vector_side:right"
 89557 .99398
 /note="assembly_fragment"
 99499 .105224
 /note="assembly_fragment"
 105325 .108843
 /note="assembly_fragment"
 108944 .112872
 /note="assembly_fragment"

ORIGIN

Alignment Scores:
 Pred. No.: 1.52e-99 Length: 112872
 Score: 1112.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.35% Indels: 0
 DB: 2 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AP000716 (1-112872)

QY 1 LysileAsnArgilleArqInValThrSerGlyGluileSerLeuValAlaGlyAla 20
 Db 77607 AAGATCAACCCCATCAGCAGGTCAACACTAGTGAGAGATCTCACTCGTGTGGGCC 77666
 QY 21 ProSerGlyCysAspCysLysAsnAlaAsnAlaAsnAlaAsnAlaAsnAlaAsnAla 40
 Db 77667 CCCAGTGGCTGTGACTGTAAATGATGCCAATCTGTGATGTTTCTGGAGAGCATGGT 77726
 QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
 Db 77727 TATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGAG 77786
 QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
 Db 77787 CTCTACGTGGCCGACCTTGGGAATCCGAATTCGGTTTATCCGGAAGAAACAGCCTTC 77846
 QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
 Db 77847 CTCACACCCAGAACATGTATGACTCTCTTCCCAATTCACCAAGAGAGCTCTATCTGTTT 77906
 QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
 Db 77907 GATACCAACCCGCAAGCACCTGTATACCCAAAGCCTGCCACAGAGACTACCTGTACAAC 77966
 QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
 Db 77967 TTCACCTACTCGGGACGCGGACATCACACTCATCACACCAACAATGGCAACATGGTA 78026
 QY 141 AsnValArgArgAspSerThrGlyMetProLeuTripleuValValProAspGlyGlnVal 160
 Db 78027 AATGTCGCGCGAGACTCTACTGGGATGCCCTCTGGCTGTGTCCAGATGGCAGGTG 78086
 QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
 Db 78087 TACTGGGTGACCATGGCGCACCAACAGTGCCTCAAGAGTGTGACCAACAGACACGAG 78146
 QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
 Db 78147 TTGGCCATGATGACATACCATGGCAATTCGGCTTCTGGCAACCAAAAGCAATGAAAC 78206
 QY 201 GlyTripThrThrPheTyrGlu 207
 Db 78207 GGATGGACAACATTTTATGAG 78227

RESULT 10
 AP001141
 LOCUS
 DEFINITION
 Homo sapiens chromosome 11 clone RP11-646020 map 11q14, WORKING
 DRAFT SEQUENCE, 23 unordered pieces.
 AP001141
 ACCESSION
 AP001141.2 GI:8118459
 VERSION
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 Homo sapiens (human)
 SOURCE
 Homo sapiens
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 165812)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens 165,812 genomic DNA of 11q14
 Published Only in Database (2000)
 2 (bases 1 to 165812)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,


```

misc_feature /note="assembly_fragment"
141855..146576
misc_feature /note="assembly_fragment"
146677..151463
misc_feature /note="assembly_fragment"
151564..154735
misc_feature /note="assembly_fragment"
154836..156703
misc_feature /note="assembly_fragment"
156804..158461
misc_feature /note="assembly_fragment"
158562..160497
misc_feature /note="assembly_fragment"
160598..162229
misc_feature /note="assembly_fragment"
162330..164116
misc_feature /note="assembly_fragment"
164217..165812
/note="assembly_fragment"

ORIGIN
Alignment Scores:
Pred. No.: 2 47e-99 Length: 165812
Score: 1112.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.35% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AP001141 (1-165812)

QY 1 LysileAsnArgIleArgGlnValThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 148342 AAGATCAACCGCATCAGCAGGTCCACACTAGTGGAGAGATCTCCTCGTTGCTGGGGCC 148401
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 148402 CCCAGTGGCTGTGACTGTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 148461
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 148462 TATGCCAGGATGCAAGATTAATATACCCATCTTCCTGTGGTGTGTGTGTGTGTGTGTGT 148521
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 148522 CTCTACGTGGCCGACCTTGGACATCCGATTCGGTTTATCCGAAGACCAAGCCCTTC 148581
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 148582 CTCACACCCAGAACATGATATGAGCTGTCTTCCCAATTGACCGAGGAGCTCTATCTGTTT 148641
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 148642 GATACACCGCGCAAGCACCTGTACCCAAAGCTGCCCAAGGAGACTACCTGTGTACAC 148701
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 148702 TTCACCTACTCGGCGAGCGGACATCACACTCATCACAGACCAACATGSCACATGTTA 148761
QY 141 AsnValArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 148762 AATGTCGCGCGAGACTCTACTGGGATGCCCTCTGGCTGTGTGTGTGTGTGTGTGTGTGTGT 148821
QY 161 TyrTyrValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 148822 TACTGGGTGACCATGGGACCAACAGTGTGACTCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 148881
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 148882 TTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAAAC 148941
QY 201 GlyTyrThrThrPheTyrGlu 207
|||||

```

Db 148942 GGATGGACACATTTATGAG 148962

RESULT 11

AP002412

LOCUS 170156 bp DNA linear HTG 31-MAY-2000

DEFINITION Homo sapiens chromosome 11 clone RP11-7H7 map 11q, WORKING DRAFT

SEQUENCE, 18 unordered pieces.

ACCESSION AP002412

VERSION AP002412.1 GI:8131676

KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 170156)

Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

Homo sapiens 170,156 genomic DNA of 11q

Published Only in Database (2000)

2 (bases 1 to 170156)

Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@psc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)

Center: RIKEN Genomic Sciences Center (GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@psc.riken.go.jp

----- Project Information

Center project name: HumDraft11

Center clone name: RP11-7H7

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 159074 bases at least Q40

Consensus quality: 164580 bases at least Q30

Consensus quality: 167092 bases at least Q20

Insert size: 168456; sum-of-contigs

Quality coverage: 5.57x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 18 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 29785 contig of 29785 bp in length

29886 52941 contig of 23056 bp in length

53042 71728 contig of 18687 bp in length

71829 87672 contig of 15844 bp in length

87773 102073 contig of 14301 bp in length

102174 109422 contig of 7249 bp in length

109523 117422 contig of 7900 bp in length

117523 124911 contig of 7389 bp in length

125012 131158 contig of 6147 bp in length

131259 135747 contig of 4489 bp in length

135848 141547 contig of 5700 bp in length

141648 147053 contig of 5406 bp in length

147154 151529 contig of 4376 bp in length

151630 156372 contig of 4743 bp in length

156473 161167 contig of 4695 bp in length

161268 164102 contig of 2835 bp in length

164203 167603 contig of 3401 bp in length

167704 170156 contig of 2453 bp in length.

* NOTE: This is a 'working draft' sequence. It currently

```

* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 29785: contig of 29785 bp in length
* 29786 29885: gap of 100 bp
* 29886 52941: contig of 23056 bp in length
* 52942 53041: gap of 100 bp
* 53042 71728: contig of 18687 bp in length
* 71729 71828: gap of 100 bp
* 71829 87672: contig of 15844 bp in length
* 87673 87772: gap of 100 bp
* 87773 102073: contig of 14301 bp in length
* 102074 102173: gap of 100 bp
* 102174 109422: contig of 7249 bp in length
* 109423 109522: gap of 100 bp
* 109523 117422: contig of 7900 bp in length
* 117423 117522: gap of 100 bp
* 117523 124911: contig of 7389 bp in length
* 124912 125011: gap of 100 bp
* 125012 131158: contig of 6147 bp in length
* 131159 131258: gap of 100 bp
* 131259 135747: contig of 4489 bp in length
* 135748 141547: gap of 100 bp
* 141548 141647: gap of 100 bp
* 141648 147053: contig of 5406 bp in length
* 147054 147153: gap of 100 bp
* 147154 151529: contig of 4376 bp in length
* 151530 151629: gap of 100 bp
* 151630 156372: contig of 4743 bp in length
* 156373 156472: gap of 100 bp
* 156473 161167: contig of 4695 bp in length
* 161168 161267: gap of 100 bp
* 161268 164102: contig of 2835 bp in length
* 164103 164203: contig of 100 bp
* 164204 167703: gap of 3401 bp in length
* 167704 170156: contig of 100 bp
* 167704 170156: contig of 2453 bp in length.
FEATURES
      source
        1..170156
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"
          /chromosome="11"
          /map="11q"
          /clone="RP11-7H7"
        1..29785
          /note="assembly_fragment"
        29886..52941
          /note="assembly_fragment"
        53042..71728
          /note="assembly_fragment"
        71829..87672
          /note="assembly_fragment"
        87773..102073
          /note="assembly_fragment"
        102174..109422
          /note="assembly_fragment"
        109523..117422
          /note="assembly_fragment"
        117523..124911
          /note="assembly_fragment clone_end:R7 vector_side:left"
        125012..131158
          /note="assembly_fragment clone_end:SP6 vector_side:left"
        131259..135747
          /note="assembly_fragment"
        135848..141547
          /note="assembly_fragment"
        141648..147053
          /note="assembly_fragment"

```

```

misc_feature
147154..151529
/note="assembly_fragment"
misc_feature
151630..156372
/note="assembly_fragment"
misc_feature
156473..161167
/note="assembly_fragment"
misc_feature
161268..164102
/note="assembly_fragment"
misc_feature
164203..167703
/note="assembly_fragment"
misc_feature
167704..170156
/note="assembly_fragment"
ORIGIN
Alignment Scores:
Pred. No.: 2.55e-99 Length: 170156
Score: 1112.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.35% Indels: 0
DB: 2 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x AP002412 (1-170156)
QY 1 LysilleAsnArgilleArgGlnValThrThrSerGlyGluilleSerleuValalaglyAla 20
Db 143299 AAGATCAACCGCATCAGGCGAGTCACCTAGTGGAGAGATCTCCTGCTGGGGCC 143358
QY 21 ProSerGlyCysAspCyslyshAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 143359 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATGTTGTTCTGGACGATGTT 143418
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 143419 TATGCCAAGGATGCAAGTTAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGAG 143478
QY 61 LeuTyrValAlaAspLeuGlyAsnilleArgilleArgPheilleArglyAsnlylsProPhe 80
Db 143479 CTCTACGTGGCGGACCTTGGGAACATCCGAATTCGGTTTATCCGAAGAACAAGCCCTTC 143538
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProilleAspGlnGluLeuTyrLeuPhe 100
Db 143539 CTCACACCCAGAACATGATGAGCTGTCTTACCAATTTGACAGGAGCTCTATCTGTTT 143598
QY 101 AspThrThrGlylyshisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 143599 GATACCCCGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAAC 143658
QY 121 PheThrTyrThrGlyAspGlyAspilleThrLeuilleThrAspAsnAsnGlyAsnMetVal 140
Db 143659 TTCACCTACACTGGGAGCGGCGACATCACACTCATCACAGACAAATGGCAACATGGTA 143718
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTripleuValValProAspGlyGlnVal 160
Db 143719 AATGTCCGCGAGACTCTACTGGGATGCCCTCTGGCTGTGTGCCAGATGCCAGGTG 143778
QY 161 TyrTripleuThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 143779 TACTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGACCAACAGGACACGAG 143838
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 143839 TTGGCCATGATGATACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAAC 143898
QY 201 GlyTripleuThrPheTyrGlu 207
Db 143899 GGATGGACAAACATTTATGAG 143919
RESULT 12
AP002768/c AP002768 186084 bp DNA linear PRI 15-MAR-2003
LOCUS
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-673F18,

```

```

complete sequence.
ACCESSION AP002768
VERSION AP002768.3 GI:13429926
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens Genomic DNA
Published Only in Database (2000)
2 (Bases 1 to 186084)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Mar 21, 2001 this sequence version replaced gi:11071944.
COMMENT Location/Qualifiers
FEATURES
    source
        1..186084
            /organism="Homo sapiens"
            /mol_type="Genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="11"
            /map="11q"
            /clone="RP11-673F18"

ORIGIN
Alignment Scores:
Pred. No.: 2,86e-99 Length: 186084
Score: 1112.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.35% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AP002768 (1-186084)

QY 1 LysileAsnArgIleArgGlnValThrSerGlyGluIleSerLeuValAlaGlyVala 20
Db AAGATCAACCGCATCAGCGAGGTACACATAGTGGAGAGATCTCACTCTGTGGGGCC 71630
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db CCAGTGGCTGTGACTGTAAATAATGATGCCAATCTGTAATGTTTCTGGAGACGATGGT 71570
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db TATGCCAAGATGCAAGTTAAATACCCCATCTCTCTGTGCTGTGTGCTGATGGGGAG 71510
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db CTCTACGTGGCGACCTTGGGAACATCCGAATTCGGTTATCCGGAGAACAGCCCTTTC 71450
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db CTCACACCCAGAACATGATGAGCTGTCTTACCAATTCACCAAGAGCTCTATCTGTTT 71390
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db GATACACCGCGCAGACCTGTACACCCAAAGCCTGCCACAGAGACTACCTGTGTACAAC 71330
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db TTCACCTACACTGGGCGCGGACATCATCATCATCATCATCATCATCATCATCATCATCAT 71320
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrValValProAspGlyGlnVal 160

```

```

Db 71269 AATGTCGCGGAGACTTACTGGGATGCCCCCTCTGGTGGTGGTCCAGATGCCAGGTG 71210
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db TACTGGGTGACCATGGGACCAACAGTGCACCTCAGAGATGTGACCAACACAGACACGAG 71150
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db TTGGCCATGATGACATACCATGGCAATTCGGCCCTCTGGCAACCAAAAGCAATGAAAC 71090
QY 201 GlyTrpThrThrPheTyrGlu 207
Db GGATGGACACATTTTATGAG 71069
RESULT 13
AB026980
LOCUS AB026980 9264 bp mRNA linear VRT 05-NOV-1999
DEFINITION Danio rerio mRNA for ten-m4, complete cds.
ACCESSION AB026980
VERSION AB026980.1 GI:5307784
KEYWORDS ten-m4.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (sites)
REFERENCE Mieda,M., Kikuchi,Y., Hirate,Y., Aoki,M. and Okamoto,H.
AUTHORS Compartmentalized expression of zebrafish ten-m3 and ten-m4,
TITLE homologues of the Drosophila ten(m)/odd Oz gene, in the central
JOURNAL nervous system
MEDLINE Mech. Dev. 87 (1-2), 223-227 (1999)
PUBMED 99425191
REFERENCE 2 (bases 1 to 9264)
AUTHORS Mieda,M.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1999) Michihiro Mieda, Brain Science Institute,
Wako-shi, Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp,
Tel:81-48-467-9713, Fax:81-48-467-9714)
COMMENT Sequence updated (29-Jun-1999).
FEATURES
    Location/Qualifiers
        1..9264
            /organism="Danio rerio"
            /mol_type="mRNA"
            /db_xref="taxon:7955"
            /dev_stage="embryo"
        1..9264
            /genes="ten-m4"
            /genes="ten-m4"
            /notes="similar to Drosophila melanogaster tenm/odz and Mus
            musculus Doc4"
            /codon_start=1
            /product="ten-m4"
            /protein_id="BAA81893.1"
            /db_xref="GI:5307785"
            /translation="MEVKRRPYRSITSRDTERRTYSSADSEDKINPKYSVSET
            LKAFQDSRLAYGSRVKDLYHEADRFQGGDFSLRDMAFGDPVPPHMGAYTEMGL
            PHRDYSVVASDADTETDGMSPHRAVLWGRNKTSGRSCLFSRANSLTLTDTFTH
            ENTENPPHCCSSSSPVDPYPHAAQNSQGRLLGNAGQAGQAGQAGQAGQAGQAGQAGQ
            SFVLKSGNVCVTPAATANECSFQNSRLRTPPLPLFHSHPQHTASIGLSRSD
            YTORNSPAPTDSPAPNEGTSADSSAQNLLNSNVPLETRNIAKTQLETLQD
            TFSRPAFSIKKPKYKCNKWKCAALSAILSVTLFLAYFIAMHFLGNHLOFQVQRI
            YOLTDENTSGHLPTDGLPLPLNGTLEFPDRGSDGKLDGFFPDSIDMGEIDVG
            RKVAQLIPGIFWRSQVFDHFMPLKFNVLSDKDALVIGRRPLPSPHQTQDFVELL
            DGRLLSQGLFGLDGPPFPFAQORSLVPTSHDTGCIQYMDSGIWHLAVYNDGKETEYQ
            SPLTATIDSIDDPCNCFGNGCHCFPGFGPDCSRASCPLVCSNGQYLKGR
            CMHSGWKSCECDVPTNOCIDITCSGHGTCIVGTICINPSYKGENCEVDLDPTCSG
            RGVCVRGCHCFVGVGGGCGSRASCMQCGSGFLADTNTCNCHNWTGDCSTG
            LCAACGGHGI CVAGSCRCDEGMGTGCEQRAChPRCSEHGTCCKDGKCECSPGNGEH

```


coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

| FEATURES | | Location/Qualifiers | |
|---------------|--|---------------------|--|
| source | 1..226256 | | |
| misc_feature | /organism="Mus musculus" | | |
| | /mol_type="genomic DNA" | | |
| | /db_xref="taxon:10090" | | |
| | /chromosome="7" | | |
| | /clone="RP23-8M3" | | |
| repeat_region | complement(147..670) | | |
| | /note="Region similar to Mm#SL670873 601761001F1 Mus musculus cDNA: Bf124644" | | |
| | 938..971 | | |
| | /rpt_family="B1-F" | | |
| | /rpt_family="B1-F" | | |
| repeat_region | 1118..1165 | | |
| | /rpt_family="T-rich" | | |
| | complement(2209..2323) | | |
| | /rpt_family="L2" | | |
| | 2468..2493 | | |
| repeat_region | /rpt_family="CAGA)n" | | |
| | complement(2497..2613) | | |
| | /rpt_family="B1-F" | | |
| | 3397..3501 | | |
| | /rpt_family="MIR" | | |
| repeat_region | 3758..3948 | | |
| | /rpt_family="B4A" | | |
| | 5668..5732 | | |
| | /rpt_family="B4" | | |
| | 5733..5809 | | |
| repeat_region | /rpt_family="PBID10" | | |
| | complement(5737..5848) | | |
| | /rpt_family="YREP_MM" | | |
| | 5810..5820 | | |
| | /rpt_family="B4" | | |
| repeat_region | 6217..6254 | | |
| | /rpt_family="(TTGG)n" | | |
| | 6509..6566 | | |
| | /rpt_family="(TG)n" | | |
| | 6721..6756 | | |
| repeat_region | /rpt_family="(CA)n" | | |
| | complement(7827..7942) | | |
| | /rpt_family="PBID10" | | |
| | complement(7899..7964) | | |
| | /rpt_family="B4" | | |
| repeat_region | complement(7965..8038) | | |
| | /rpt_family="ID4" | | |
| | 8093..8191 | | |
| | /rpt_family="L2" | | |
| | 8326..8437 | | |
| repeat_region | /rpt_family="MIR" | | |
| | 9579..9616 | | |
| | /rpt_family="(CATA)n" | | |
| | 9641..9664 | | |
| | /rpt_family="(CCA)n" | | |
| repeat_region | complement(9715..9783) | | |
| | /rpt_family="PBID10" | | |
| | complement(9778..9830) | | |
| | /rpt_family="ID3" | | |
| | 9927..9958 | | |
| repeat_region | /rpt_family="(CAAA)n" | | |
| | complement(10121..10237) | | |
| repeat_region | /rpt_family="B1-F" | | |
| | 10243..10414 | | |
| | /rpt_family="(CCTG)n" | | |
| | 10440..10509 | | |
| | /rpt_family="CT-rich" | | |
| repeat_region | complement(10532..10578) | | |
| | /rpt_family="B1-F" | | |
| | 10971..11355 | | |
| | /rpt_family="Lx6" | | |
| | complement(11571..13187) | | |
| misc_feature | /note="Region similar to Mm#S685488 Odd Oz/ten-m homolog 4 (Drosophila): AB025413" | | |
| | 14505..14614 | | |
| | /rpt_family="PB1" | | |
| | 15111..15194 | | |
| | /rpt_family="MIR" | | |
| repeat_region | complement(15863..16043) | | |
| | /rpt_family="B4" | | |
| | 16385..16459 | | |
| | /rpt_family="MIR" | | |
| | 19475..19693 | | |
| repeat_region | /rpt_family="B2" | | |
| | 20006..20169 | | |
| | /rpt_family="B3" | | |
| | 20170..20190 | | |
| | /rpt_family="(TA)n" | | |
| repeat_region | 20191..20251 | | |
| | /rpt_family="B3" | | |
| | complement(21402..21484) | | |
| | /rpt_family="MIR" | | |
| | 24484..24563 | | |
| repeat_region | /rpt_family="L2" | | |
| | 24789..24907 | | |
| | /rpt_family="L2" | | |
| | complement(25094..25241) | | |
| | /rpt_family="MIR" | | |
| repeat_region | complement(26937..27347) | | |
| | /rpt_family="L1MCC" | | |
| | 27569..27613 | | |
| | /rpt_family="(TTTA)n" | | |
| | complement(28594..28633) | | |
| repeat_region | /rpt_family="L2" | | |
| | complement(28841..28955) | | |
| | /rpt_family="PBID10" | | |
| | 29428..29480 | | |
| | /rpt_family="(TATG)n" | | |
| repeat_region | 29510..29582 | | |
| | /rpt_family="(CAGA)n" | | |
| | 29583..29723 | | |
| | /rpt_family="GA-rich" | | |
| | 29776..29959 | | |
| repeat_region | /rpt_family="(CAGAGA)n" | | |
| | complement(30193..30338) | | |
| | /rpt_family="B1_MM" | | |
| | 31162..31452 | | |
| | /rpt_family="B4" | | |
| repeat_region | 31913..31950 | | |
| | /rpt_family="AT-rich" | | |
| | 32864..32885 | | |
| | /rpt_family="AT-rich" | | |
| | complement(33023..33070) | | |
| repeat_region | /rpt_family="MIR" | | |
| | 34393..34458 | | |
| | /rpt_family="MIR" | | |
| | complement(34687..34859) | | |
| | /rpt_family="MER5B" | | |
| repeat_region | 34872..34950 | | |
| | /rpt_family="MER5A" | | |
| | 35660..35869 | | |
| | /rpt_family="URR1A" | | |
| | 36089..36106 | | |
| repeat_region | /rpt_family="B4" | | |
| | 36107..36239 | | |

```

/rpt_family="BI_MM"
36263..36295
/rpt_family="(CCA)n"
36398..36419
/rpt_family="(CCCCAG)n"
36790..36817
/rpt_family="(T)n"
complement(38755..38926)
/rpt_family="B4"
complement(39294..39403)
/rpt_family="L2"
40007..40038

Alignment Scores:
Pred. No.: 9,298-98 Length: 226256
Score: 1098.00 Matches: 204
Percent Similarity: 99.52% Conservative: 2
Best Local Similarity: 98.55% Mismatches: 1
Query Match: 79.34% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AC073599 (1-226256)

QY 1 LysileAsnArgileArgGlnValThrThrSerGlyGluileSerLeuValAlaGlyAla 20
Db 33926 AAGATCAATCGCATCAGGCAGGTCCTACTACAGTGGTGAGATCTCACTGCTTGTGTGTC 33867

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 33866 CCCAGTGGCTGTGACTGTGTAATAATGATGCAACTGTGCTGCTCTCTCGAGATGATGT 33807

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 33806 TAGCCCAAGGATGCAAGCTGAATACCCCATGCTCTGCTGTGTGTGCTGACGGGGAG 33747

QY 61 LeuTyrValAlaAspLeuGlyAsnileArgileArgPheileArgLysAsnLysProPhe 80
Db 33746 CTTCTATGTGGCCGACCTGGGAACATCCGATTCGATTATCCGGAAGAACAGCTTTC 33687

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProileAspGlnGluileuTyrLeuPhe 100
Db 33686 CTGAACACTCAGAACATGTACGAGCTATCTCTCCCATCGACAGAGCTGTACCTCTTT 33627

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 33626 GATACCAAGTGGCAAGCATCTGTACATCAGAGCTACCTACAGGGGACTACCTGTACAC 33567

QY 121 PheThrTyrThrGlyAspGlyAspilleThrLeuileThrAspAsnAsnGlyAsnMetVal 140
Db 33566 TTCCTTACACAGGGGAGCGGACATCACACATATCACCGAACAAATGGCAACATGGTG 33507

QY 141 AsnValArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
Db 33506 AACCTCGCGGAGACTCTACCGGATCGCTCTCTGCTGTGTGTGTGCTGAGATGCGGTA 33447

QY 161 TyrTyrValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 33446 TACTGGGTACCAATGGGACCAACAGCGGCTCAGAGTGTGACCAACAGACACGAG 33387

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 33386 CTAGCCATGATGACCTACCATGCGCACTCTGCGCTCTTGTGGCAACCAAAAGCAATGAAAC 33327

QY 201 GlyTyrThrThrPheTyrGlu 207
Db 33326 GGGTGGACACGTTTATGAG 33306

```

RESULT 15
AC120288/c AC120288 245722 bp DNA linear HTG 15-NOV-2002
LOCUS Rattus norvegicus clone CH230-17M17, WORKING DRAFT SEQUENCE.
DEFINITION AC120288
AC120288
AC120288.4 GI:25008439
VERSION

KEYWORDS SOURCE ORGANISM

HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE AUTHORS

1 (bases 1 to 245722)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, Y.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinu, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunarathne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, Y., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J.,
Lorensu, H., L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Manum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milesavljic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Moran, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, W., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

TITLE JOURNAL

Direct Submission
Unpublished
2 (bases 1 to 245722)

REFERENCE AUTHORS

Direct Submission
Submitted (05-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

TITLE JOURNAL

3 (bases 1 to 245722)
Rat Genome Sequencing Consortium.

REFERENCE AUTHORS

Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

TITLE JOURNAL

On Nov 15, 2002 this sequence version replaced gi:23669170.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas

(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWUK
Center clone name: CH230-17M17

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 233443 bases at least Q40
Consensus quality: 234913 bases at least Q30
Consensus quality: 236024 bases at least Q20
Estimated insert size: 236455; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 245722: contig of 245722 bp in length.

FEATURES

```

source
1..245722
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-17M17"
1..1076
/note="wgs end extension
clone_end:Sp6"

misc_feature
6381..7227
/note="clone_boundary
clone_end:Sp6
site:
end sequence:BH280179"
6388..6810
/note="clone_boundary
clone_end:Sp6
site:
end sequence:BH280179"
complement(73338..240671)
/note="clone_boundary
clone_end:T7
site:
end sequence:BH280178"
complement(240061..240644)
/note="clone_boundary
clone_end:T7
site:
end sequence:BH280178"
244240..245722
/note="wgs end extension
clone_end:Sp6"

```

ORIGIN

Alignment Scores:

| Pred. No.: | 5.19e-97 | Length: | 245722 |
|------------------------|----------|---------------|--------|
| Score: | 1091.00 | Matches: | 203 |
| Percent Similarity: | 99.03% | Conservative: | 2 |
| Best Local Similarity: | 98.07% | Mismatches: | 2 |
| Query Match: | 78.83% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-10-029-020-14_COPY_1490_1750 (1-261) x AC120288 (1-245722)

| | | | |
|----|-------|---|-------|
| QY | 1 | LysilleAsnArgileArgGlnValThrThrSerGlyGluIleSerIleValAlaGlyAla | 20 |
| Db | 15709 | AAATCAACCGTATCAGCAGGTACGACAGTGTGAGATCTCACTGCTTCTGCTGTC | 15650 |
| QY | 21 | ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly | 40 |
| Db | 15649 | CCCAGCGCTGCGACTGTAAAAACGACGCCAACTCGGACTGCTTCTCTGGAGATGCGC | 15590 |
| QY | 41 | TyrAlaLysAspAlaLysIleuAsnThrProSerSerLeuAlaValCysAlaAspGly | 60 |
| Db | 15589 | TACGCCAAGGACGCAAGCTGAACACCCCATCTTCTTGGCCGTGTGTGACACGGGAG | 15530 |
| QY | 61 | LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe | 80 |
| Db | 15529 | CTCTACGTGGCGGACCTTGGAAACATCCGATCCGTTTATCCGGAGAACAAAGCTGTC | 15470 |
| QY | 81 | LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe | 100 |
| Db | 15469 | CTGAACACCCAGAACATGTATGAGCTGTCTCTCCCATCGACAGGAGCTGTACTCTTT | 15410 |
| QY | 101 | AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn | 120 |
| Db | 15409 | GACACCAAGTGGCAAGCATCTGTACACTCAGAGCCTGCCACAGGAGACTACCTGTACAAC | 15350 |
| QY | 121 | PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal | 140 |
| Db | 15349 | TTCATTACAGCGGGGACGGGACATCACACATCATCAGACAAACAAATGGCAATGGTG | 15290 |
| QY | 141 | AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal | 160 |
| Db | 15289 | AACGTCCGCGCAGATTCTACCGGATGCCTCTCTGCTGTGTAGTCCAGATGGCAGGTG | 15230 |
| QY | 161 | TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu | 180 |
| Db | 15229 | TACTGGGTAAACCATGGGCACCAACAGCGCACTCAGAAAGTGTGACACGACGACACGAG | 15170 |
| QY | 181 | LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn | 200 |
| Db | 15169 | CTGGCTATGATGACATACCATGGCACTCTGGCTCTTGGCGACCAAAAGCAATGAAAC | 15110 |
| QY | 201 | GlyTyrThrThrPheTyrGlu | 207 |
| Db | 15109 | GGGTGGACAACGTTTTATGAG | 15089 |

Search completed: August 14, 2004, 12:00:37
Job time : 5091.26 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 1786.37 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_1100_1200

Perfect score: 549

Sequence: 1 RLFRKWFAPLDSYFIWD.....IQSGILHKNGENQFVSQOP 101

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0*

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_19065 @runat_06082004_112216_29275 -NCPV=6 -ICPV=3
-NO_MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.in:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.nam:*
37: em.htg.vrt:*
38: em.sv:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | | | DB | ID | Description |
|------------|-------|-------|--------|----|-----------|-------------|
| | Score | Match | Length | | | |
| 1 | 549 | 100.0 | 8354 | 6 | AX556500 | Sequence |
| 2 | 549 | 100.0 | 8645 | 6 | AX600210 | Sequence |
| 3 | 544 | 99.1 | 9722 | 10 | AF059485 | Mus muscu |
| 4 | 540 | 98.4 | 8585 | 10 | AB025413 | Mus muscu |
| 5 | 538.5 | 98.1 | 8438 | 6 | AX675551 | Sequence |
| 6 | 484 | 88.2 | 9264 | 5 | AB026980 | Danio rer |
| 7 | 454 | 82.7 | 5804 | 10 | AK122513 | Mus muscu |
| 8 | 454 | 82.7 | 7816 | 10 | AF195418 | Mus muscu |
| 9 | 454 | 82.7 | 8964 | 10 | AB025412 | Mus muscu |
| 10 | 453 | 82.5 | 3394 | 9 | AK125869 | Homo sapi |
| 11 | 453 | 82.5 | 5309 | 9 | AB040888 | Homo sapi |
| 12 | 453 | 82.5 | 8473 | 6 | AX662357 | Sequence |
| 13 | 453 | 82.5 | 8487 | 6 | AX662359 | Sequence |
| 14 | 453 | 82.5 | 8645 | 6 | AX662355 | Sequence |
| 15 | 453 | 82.5 | 8675 | 6 | AX662353 | Sequence |
| 16 | 446 | 81.2 | 8816 | 5 | AB026979 | Danio rer |
| 17 | 426 | 77.6 | 8409 | 5 | GGA279031 | Gallus ga |
| 18 | 426 | 77.6 | 8409 | 5 | GGA279031 | Gallus ga |
| 19 | 420 | 76.5 | 8118 | 5 | GGA238613 | Sequence |
| 20 | 417 | 76.0 | 8797 | 6 | AX250066 | Sequence |
| 21 | 417 | 76.0 | 8797 | 6 | AX250066 | Sequence |
| 22 | 417 | 76.0 | 8797 | 10 | AB025411 | Mus muscu |
| 23 | 416 | 75.8 | 6560 | 6 | AX250065 | Sequence |
| 24 | 416 | 75.8 | 7713 | 9 | HSM808325 | Homo sapi |
| 25 | 416 | 75.8 | 7781 | 9 | AB032953 | Homo sapi |
| 26 | 411 | 74.9 | 8297 | 9 | AF100772 | Homo sapi |
| 27 | 408 | 74.3 | 8373 | 10 | AB025410 | Mus muscu |
| 28 | 407 | 74.1 | 9729 | 6 | AX250013 | Sequence |
| 29 | 407 | 74.1 | 9826 | 6 | AX250008 | Sequence |
| 30 | 406 | 74.0 | 8689 | 6 | AX250067 | Sequence |
| 31 | 406 | 74.0 | 8689 | 10 | AF086607 | Rattus no |
| 32 | 323 | 58.8 | 8575 | 6 | AX921803 | Sequence |
| 33 | 265 | 48.3 | 112872 | 2 | AF000716 | Homo sapi |
| C 34 | 265 | 48.3 | 170156 | 2 | AF002412 | Homo sapi |
| C 35 | 265 | 48.3 | 186084 | 9 | AP002768 | Homo sapi |
| C 36 | 261 | 47.5 | 226256 | 10 | AC073599 | Mus muscu |
| C 37 | 261 | 47.5 | 245722 | 2 | AC120288 | Rattus no |
| C 38 | 261 | 47.5 | 257693 | 2 | AC103323 | Rattus no |
| C 39 | 260 | 47.4 | 180659 | 5 | EX001038 | Zebrafish |
| 40 | 248.5 | 45.3 | 10471 | 3 | AF008228 | Drosophil |
| 41 | 248.5 | 45.3 | 11451 | 3 | DMTENN | Drosophil |
| 42 | 248.5 | 45.3 | 14045 | 3 | DMOD24 | Drosophil |
| 43 | 248.5 | 45.3 | 82852 | 2 | AC017685 | Drosophil |
| C 44 | 248.5 | 45.3 | 169841 | 3 | AC009366 | Drosophil |
| C 45 | 248.5 | 45.3 | 305612 | 3 | AB003597 | Drosophil |

ALIGNMENTS

RESULT 1

| | |
|--|--|
| Dugan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K., Khare,R. and Walla,N.K. | |
| Proteins associated with cell growth, differentiation, and death | |
| Patent: WO 02072830-A 22 19-SEP-2002; | |
| Incyte Genomics, Inc. (US) | |
| FEATURES | Location/Qualifiers |
| source | 1..8645 |
| | /organism="Homo sapiens" |
| | /mol_type="unassigned DNA" |
| | /db_xref="taxon:9606" |
| ORIGIN | /note="Incyte ID No: 7488573CB1" |
| Alignment Scores: | |
| Pred. No.: | 1,25e-58 |
| Score: | 549.00 |
| Percent Similarity: | 100.00% |
| Best Local Similarity: | 100.00% |
| Query Match: | 100.00% |
| DB: | 6 |
| US-10-029-020-14_COPY_1100_1200 (1-101) x AX600210 (1-8645) | |
| Qy | 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTrpPheIleTrpAsp 20 |
| Db | 3402 CGCCTCTTCAGGAAGTGGTTCGTCGAGCCCGACAGCTGTCCTATTATTTTCATTTGGGAC 3461 |
| Qy | 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40 |
| Db | 3462 AAGACAGACGCTACCAACCAAGAGGTGTTTGGGCTTTCAGAGCCCTTGTTCGTGGGT 3521 |
| Qy | 41 TyrGluTyrGluSerCysProAspLeuLeuLeuTyrGluLysArgThrValLeuGln 60 |
| Db | 3522 TATGATATGAAATCCTGCCAGATCAATCTCTGGGAAAAAGAACACAGTGTCTGCAG 3581 |
| Qy | 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80 |
| Db | 3582 GGCCTATGAAATTCATGCGTCCAAAGCTTGGAGGATGGAGCCTAGACAAACATCATGCCCTC 3641 |
| Qy | 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100 |
| Db | 3642 AACATTCAAAGTGGCATCTCTGCACAAAGGAATGGGAGAACCAAGTTTGTCCTCAGCAG 3701 |
| Qy | 101 Pro 101 |
| Db | 3702 CCT 3704 |
| RESULT 3 | |
| AF059485 | 9722 bp mRNA linear ROD 15-AUG-1998 |
| LOCUS | Mus musculus DOC4 (Doc4) mRNA, complete cds. |
| DEFINITION | |
| ACCESSION | AF059485 |
| VERSION | AF059485.1 GI:3170614 |
| KEYWORDS | |
| SOURCE | Mus musculus (house mouse) |
| ORGANISM | Mus musculus |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 9722) |
| AUTHORS | Wang,X.-Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P., Zinszner,H. and Ron,D. |
| TITLE | Identification of novel stress-induced genes downstream of chop |
| JOURNAL | EMBO J. 17 (13), 3619-3630 (1998) |
| MEDLINE | 98315054 |
| PUBMED | 9649432 |
| REFERENCE | 2 (bases 1 to 9722) |
| AUTHORS | Wang,X.-Z. and Ron,D. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (14-APR-1998) Skirball Institute, New York University |
| FEATURES | Medical Center, 550 First Ave., New York, NY 10016, USA |
| source | Location/Qualifiers |
| | 1..9722 |
| | /organism="Mus musculus" |

```
/mol_type="mRNA"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/cell_line="NIH-3T3"
/note="tunicamycin-treated"
1..9722
gene
/genes="Doc4"
/note="downstream of CHOP 4"
583..9060
CDS
/genes="Doc4"
/note="similar to Drosophila melanogaster temm/odz and human gamma-heretulin; type II transmembrane protein"
/codon_start=1
/product="DOC4"
/protein_id="AAC31807.1"
/db_xref="GI:3170615"
/translation="MDVKERKPYRLRRDAERRYTSSADSEBKGPKQSYSSSET
LKAYDQDARLAYSRVKOMVPEAEFEFCRTGNTFLRELGLCEMTPPHGTLYRTDGL
PHCYGSMGASDADLEADTVLSPHPVRLWGRSTRSGRSCLSRPNANLTLDTHEH
NTEGAPLHCSASDSTPEQSPSPSPANESORRLGNGVAQTPPSDSEEBEPVN
SELVKSASPLGVAANDHPSLQHPRLRTPPLPHAHNPQHHAASINLNRGNT
PRSPSPRTHDSLGEPPAGAQEPHAQDNWVLSKIPVETNRLGQPPFLPQWQDN
LIEMDIFASRDGAYSDEPPHFFPGGTSPLFTCTSPGPISTVSPPPRPLPRST
FSRFAFNKPKFSKYCNMKCAALSAILI SATVILILAYFVAMEHLFLNWHLOPMEGOMQ
MYEITDASSWPVPTDVSYPGSGTLETPDKGKAEGKPSLSPEDSPIDSGEI
DVGRASOKIPGTFWRSQVFLDHPVHLKENVSLGKAALVGIYRGKGLPPSHQLDFV
ELLDRRLLTAEBSLEGPQRSGRPVPPSHETGFIQYLDOSGIWHIAFYNDGKSEV
VSFLITALESDNCPNSGNGDCISGTCFLGLPGLDGRASCPVLCSGNGQYMKG
RCLCHWGKAECIVPTNQDIAVACSHGTCIMGTICNPGYKGESCEBVDMPDCTS
SRGVGREGCHSVGWGTCNETPRATCLDQSGHGTPLDPTGLCNDPSPMTGDCSI
EICAAAGCGVCGVGRCTLDLNGHVCQLGWRGCTGDSNETCGDCKGCECTPGWGE
HCTIEGPGLCNGRGCTLDLNGHVCQLGWRGCTGDSNETCGDCKGCECTPGWGE
CMPDCCLOPLCHVNPCLCLSPDLIDIIQETQAPVQSNLNSFYDRIKFLVNDGSDHS
IPENPDGGHACVNPQMTSDGTLFVPMETIVMRHEENIEPSRDLNSFARNPVPSP
ISILREPERAPITOEHTLWLPDFVMEITIVMRHEENIEPSRDLNSFARNPVPSP
SPLTSFASCAEKEIPEIQAQEEIIVAGCKMRLSYLSRTPGYKSVVRISLTHPT
IPENLMVHLVAMEVEGRLEKFWFAAPLDSYFTWDTKTVNQVGLSRAFAVSGVE
YESCPDLILWEKRTAVLQGLVIDASKLGSWLDKHALNQSGLIHKNGENQVFSQO
PVGIGTIMGNCRSISCPSCNGLADGNKLLAPVALTCGSDSLYGVGFNFIRFIPS
GNVTNILEMHSHPAKYILATDMSGAVFLSDTNSRRVFKVSTTVVVKVNSVEVA
GTGQCPDFDTRCGGKATEATLNPRTGIDVDFGLIYFVDGTMIRRVQNGIIST
LLGSDNLTARPLSCDSVMEISQVRLWPTDILAINPMDNSLYLDNNVLIQISNHOV
RIVAGRPMHQVPGPIDQFLLSKVAIHATLESATLAVSHNGVLVIAETDEKKNIRIQ
VTTSGEISLAVAGPSGCDKNDANCDCEGSDGYADKALNTSSLAVCADGELYAD
LGNIRIFIRKNKPELTONMYELSSPIDQELIPLDTSKGHLTYQSLTDGILYNTFY
TGQDTHITDNGNMVNRDSTMELVLPVDPGVYVWVTGNSALRSVTTQGHLE
AMTYHGNSSLTHLKSNGENWTFYEYDSFGRLTNFTPTGQVSSFRSDTSVSHOV
ETSKDDVTITLHSGSAGFYLLQDQVRNSYIYIGADGSLRLLLANGMEVALQTEPHL
LAGTVNPTQKRVNITLIDNGLNLEVRQKQARGQVTFGLRLRVHNRNLLSLDFD
RVTRTEKIYDHRKFTLRIYDOAGRPSFWSPSRLNGVNTVPSGSHIAGIQRTMS
ERVEYDQAGRITRSIFADGKMSVYLEKSWLHLSORQYIIEFDKNDRLSSVTMPN
VARCTLETIBSVGYRNIYQPPNGNASVIQDTEGHLHFTFGTRGVVYIKGLS
KLAEITYDTTKVSTDETAGMLKTVNLQEGFTCTIRYQIGLIDRQIRFTEEGM
VNARFDYNDNSFRVTSQMVAINETPLPIDLYRYDVSQKTEQGFYIYDINOII
TAVMTSHSFHDYAKMPEYQYETFRSLMYMTVOYDNGRVLKVKELGVAPYANTTRY
SYEDADQLOQTYSINDKPLVYSYDLNGLHLLSPNSARLTPLVLDLDRTRIGD
VOYKDDGSLRORGVDVFNYSAGLLIKAYNPASGWSRYRDLGRRVSSKSSH
HLOFFFYADLTNPYVHLVSHSSSEITSLYDLQHLFAMELSSGDEFYIACDNIPT
LAVFSGTGLMKIQLIYAYGEIYNPNFQIIIGYHGGLYDPLTKLVHNRGRDYDVL
AGNVSFDPHMLKRLSSNIVPFLHYMKNNNPISNSQDICEPMTDVNSMLTFFGQL
HNVIPGYKPDPTDAMEPSYELVHTQMTQBNWKSILGVQCEVQKQKAFVTLERFD
QLYGSTITSCQAPETKTKFASGSPFGKGVKPKALKDRVTVDIIISVANEDEGRFIAL
NNAHYLNHLHFTIDGVTHVFVKGPSGEGDLAILGLSGGBRTLENGVNTVVSQINTWL
SGRTTRYTDILOVYALCLNTRYGVTVDEKRVLELAKQVRAVRQVAREQQQLREGE
EGLRAWDEGKQOVLNTRGVQYDGFVTSVEQYPELSDSANNIHFMRQSEMGR"
1813..1881
misc_feature
/genes="Doc4"
/note="putative; transmembrane-region site"
ORIGIN
Alignment Scores:
Pred. No.: 9722
Score: 544.00 Matches: 100 Length: 9722
```

LKAYDQADARLAYSGRVDMVQPEAEFCRTCTNTFTRELGLGEMTPPHGTLTYRDIIGL
PHCGTSMGASSADLADLTVALSPHPVPLWGRSTESGRSSCLSRANSNLTLTDEHE
NTEYDHPSSLOQHPRLTPPPPLPHANTPQHHASINSNRGNTFRSPNSPAPTDH
SLSGEPFAGSQEPHQAQNNLLNSNPLETHNAGKQPLGLTQNLNELLMLLSASRH
DYXSDGHFLFKPGTSPFLCTSTGYPLSTSTVYSPPLPRSTFRPAPNLTAKPS
KYNWKAALSAIISATVILLAYFVAMHLFGLNHLQPMQMGOMYETITEDTASSW
PVPTDVSYPGSGTGLETPDRKGGAARPKSPSPEDSFIDSGEIDVGRRASQKIPP
GTFWRSQVFIHPVHLKFNVSLSGKAAALVIGYGRKGLPSPSTQDFVELLDGRLLTQIE
ARSURGPQRGRGPVPPSPSHETGFIQYLDSDGHWHLAFYNDKESVSWFLATVLESVD
NCPNSYNGNDCISGTCFLGFLGDCGRASCPVLCSGNGQIMGRCLCHSGWKGA
CDVPTNQICIDVACSHGTCIMGTICNFGYKSCBCEVDMDPTCSSRVCVRGECHE
SVWGTCNCTPRATCQCSGHGFLPDTGLCNCDPSWTGDCSIEICAAACGGHGV
KAGCGPCNGRCITLDLNGHWCVCOLGWRGTGCDTSMETGCGDKDNDGLDVCMD
PDCCLQHPNPLCGSPGDPDLDIILQETQAPVUSQONLAPFYDRIKFLVGRDSTHIGP
ENPFQGHACVIRGQVMTSDPPLVGNVIFINNEPLFGTISRQOGSEDLVNGGSI
ILFRAPAFITQETHTLWLPWFDFMYETIMRHEENIPSCDLSNFPARNPVSPSPPL
TSPASSCAEKGPIVEIOALQOEIIVAGCKMLSYLSRTPGYKSLRISLTHPTIPF
NLMKVLMAVGRVFRKFWFAAAPOLSVYFIWKTDVVNVKVFSGSEAFVSQYEGYES
CPDILLWKEKTRAVLOGEYDASKIGWSLKHHLNIOGILHKGNGNOFVSQOPPV
IGSILWNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGLVGVDFNYIRTFPSCNV
TNILEMRNDRHSHSPAHXYLATDPSGAVPLSDTNSRRVFKVKSITVVKDLVKN
EVVAGTQDCLFPDTRCGDGGKATEATLTPRGITVDKFGLIYFVDMIRRVQNG
IISLIGSNDLTASRPLSCDSVMEISQVLEWPTDLAINPMDSNLYLDNNVQLISE
NHQVIRVAGRPMHCOVPGIDHLLSKVAIHTALEATALAVSHNGVLYIAETDEKIN
RIQVTTSEISLVAGAPSGCDKANDCDFGDDGYAKDAKLNTPSLVACADGEL
YVADLGNIRIKIRKPKFNQNTOMYELSPDIDQELXLDTSQKHLVLTQSLPTGVLY
NFTYTGDDGDIHTIDNNGMNVNRDSTGNLWLVDPDQVTVWMTGNTSARSVTTQ
GHELAMTYHNGSLGATKSNENGWTFYFSDSFGRLNTVPTFGVSSFRSDTSV
HVQVTSXDDVTITNLASGAFYLLQDOVRNSYIIGADGSLRLLLANGMEVALQT
EPHLLAGTNPVTKRNVTLPIDGLNLVWVRQKQGVQVTFGRRLRVHNRNLLS
LQFDRVTRTEKLYDDHKEFTLRLYDQAGRPSPSLWSSRLNGVNVTPSGHLAGTOR
GEMEREDQAGRITSLRIPADGKMSYTYLEKSMVLHLHSORQYIFEDKNDRLSVY
TNPVVARQLETIRSVGYRNIYQPEGNASVIOQDETCHLLHLLTYLGTGRGVYKY
GKLSKLAETLYDTVSTYSYDEYATMLKTNVQNGEFTCTIRYRQGLDRLDRIT
EENWVARDNDYNSFRVTSMOAVINETPLPIDLYRYDDVSGKTEFGQKFGVYVDI
NOITTVMTHTKFDAYGRMEVOYEIFRSLMYMTVQVDMNMRVKKELKVPYVDI
TRYSEYDADGLOTVSINDKPLWYSYDLNGLNHLSPGNSARLTPRLYDLDRIT
RUGDVOYKMDGDFLQORGDPVPEYNSGILLIKAYNRSQWSVRXYVDGLGRVSSKS
SHSHLOFFVADUTNPKVTHLYNHSSEITSLYLOLQHLFAMELSSGDEFYIACDN
IGTFLVFGSGTGMKQIILYAYGEIYMDTNPFIIGYHGLXLYDPLTKVNHMRD
YDLVLAGRTSPDHLMKRLSSNIVPFLMYFNKNPIINSODIKCFMTDVLNWSLITF
GOLNHYIPGPDPDAMEPSYELVHTQMTQEDWNSKSLIGVCEQVOKLAFYTL
ERFDQYVSTISCOQAPETKFASSSGIFGKGVKALKDGRVTTOISVANDGRRRI
AALINNAHYLENLHFTIDGVDHYFVKPSPSEGLALIGLSGRRRLTENGWVTVSQI
NTMWSGRTRYTDIQLQYRALCNTRYITGVDEKRYVLELARQRAVRQWAEQORL
REGEGLRAWDTGKEKQVLTNRGVQYDGPFPVTSVQYPELSDSANNHIMRSEMR
R"

ORIGIN
Alignment Scores: 1.69e-57 Length: 8585
Pred. No.: 540.00 Matches: 99
Score: 98.02% Conservative: 0
Percent Similarity: 98.02% Mismatches: 2
Best Local Similarity: 98.36% Indels: 0
Query Match: 10 Gaps: 0
DB:
US-10-029-020-14_COPY_1100_1200 (1-101) x AB025413 (1-8585)
Qy 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db CGGCTCTTCAGGAAGTGGTTTGGCGAGCTCCTGACCTGTCTTATTACTTCTATCTGGGAC 3550
Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db AAGACAGATGCTCAACACCAAGAGGTGTTCGATTTCTCAGAAGCCTTTGTTTCCGTGGGT 3610
Qy 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrValLeuGln 60
Db TATGAGTATGATCTCTGCCAGATCTGATCTCTGTGGGAGAAAGACACAGCGGTCTGCAG 3670
Qy 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80

Db 3671 GGCTATGAAATGATGCTTCAAGCTGGGGGCTGGAGCTCTGGAATAGCACATGCCCTG 3730
Qy 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3731 AACATCCAGATGGCATCTCTGCACAAAGGGAATGGAGAGAACCACTTTTGTGTCCAGCAG 3790
Qy 101 Pro 101
Db 3791 CCA 3793
RESULT 5
AX675551 8438 bp DNA linear PAT 27-MAR-2003
LOCUS Sequence 1 from Patent WO02055704.
DEFINITION ACCESSION AX675551
VERSION AX675551.1 GI:29333552
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Padigaru, M., Li, L., Zerhusen, B.D., Casman, S.J., Shenoy, S.,
Spytek, K.A., Zhong, M., Gangolli, E.A., Burgees, C.E., Patturajan, M.,
Vernet, C.A., Taylor, S., Tchernev, V.T., Miller, C.E., Guo, X.,
Boldog, P.L., Grosse, W.M., Alsobrook, J.P., Gerlach, V.,
Edingermark, S., Rotherberg, M.E., Ellerman, K., Macdougall, J.,
Malvankar, U., Millet, I., Peyman, J., Smithson, G., Gunther, E. and
Stone, D.J.
TITLE Proteins, polynucleotides encoding them and methods of using the
same
JOURNAL Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)
FEATURES
source 1. 8438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 2.57e-57 Length: 8438
Pred. No.: 538.50 Matches: 101
Score: 99.02% Conservative: 0
Percent Similarity: 99.02% Mismatches: 0
Best Local Similarity: 98.09% Indels: 1
Query Match: 6 Gaps: 1
DB:
US-10-029-020-14_COPY_1100_1200 (1-101) x AX675551 (1-8438)
Qy 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db CGGCTCTTCAGGAAGTGGTTTCGTGCGAGCCAGACCTGTCTTATTATTTCATTTGGGAC 3370
Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db AAGACAGACCTCTACAAACCAAGAGGTGTTTGGGCTTTTCAGAAGCCTTTGTTTCCGTGGGT 3430
Qy 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrValLeuGln 60
Db TATGAAATGAAATCTCTGCCAGATCTATCTCTGTGGGAGAAAGACCAACAGCTGCTGCAG 3490
Qy 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db GGCTATGAAATGACGGTCCAGCTTGGAGGATGGAGCCCTAGACAAACATCATGCCCTC 3550
Qy 81 AsnIleGlnSer---GlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGln 99
Db AACATTCAAAGTGGTGGCATCTCTGCACAAAGGGAATGGGAGAACCACTTTGTGTCTCAG 3610
Qy 100 GlnPro 101
|||||

Db 3611 CAGCCT 3616

RESULT 6
AB026980
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AB026980 9264 bp mRNA linear VRT 05-NOV-1999
Danio rerio mRNA for ten-m4, complete cds.
AB026980
AB026980.1 GI:5307784
ten-m4.
Danio rerio (zebrafish)
Danio rerio

REFERENCE
AUTHORS
TITLE

1 (sites)
Mieda,M., Kikuchi,Y., Hirate,Y., Aoki,M. and Okamoto,H.
Compartmentalized expression of zebrafish ten-m3 and ten-m4,
homologues of the Drosophila ten(m)/odd Oz gene, in the central
nervous system
Mech. Dev. 87 (1-2), 223-227 (1999)

10495292
99425191
2 (bases 1 to 9264)
Mieda,M.

Direct Submission
Submitted (06-MAY-1999) Michihiro Mieda, Brain Science Institute,
RIKEN, Lab. for Developmental Gene Regulation; 2-1 Hirosawa,
Wako-shi, Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp,
Tel:81-48-467-9713, Fax:81-48-467-9714)
Sequence updated (29-Jun-1999).
Location/Qualifiers
1. .9264
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/dev_stage="embryo"
1. .9264
/gene="ten-m4"
323. .8797
/gene="ten-m4"
/note="similar to Drosophila melanogaster tenm/odx and Mus
musculus Doc4"
/codon_start=1
/product="ten-m4"
/protein_id="BA081893.1"
/db_xref="GI:5307785"
/translation="MEVKRRPRLSRRDTRRTYSSADSDGKINPKSYSSST
LKAFDQSRUAYGSRVKDLVHHEAFSRQDPDFSLRDMAGDPVPPHMGAYRTMGL
PHRDYSVASADADTETDGMSPHVAHLWGRNKGSRGSCILFSPRANSNLTDTGHL
ENTENPPPLHCSASSPVDSPPYPPSHAAQSQGRLLGNSGAQAGRDSSEDEFGN
SFLVKTSGNVCTPAATANEFSQNHSLRTPPLPLPHSHSPSHQHTASIGLSRDN
YTORSPAPPTDSSAPNEGPTSAQSSAQDNWLLNSVPLETRNIAKOTFLETQD
NFTEMILLATARDGATYDGHFLFKPGTSPLYCTTSPGYLTSSTVYSPPPRLPN
YQUTENDTSLHPLDGLPLNGTGLEFPDRSGDDKLDGFPEDSFIDMGEIDVG
RKVALIPGIFRWSQVFIIDHMYLKFNVSLSKDALVGIYGRRLPSPHQDFVEL
DGRLLSQGLDLPDPFPAQORSVPEITSHDTGCIQYMDSGIWHLAVNDKGTQV
SFLTTAIDSDDCPSNCFNGDCVSGNCHCFPGFRGDCSRACPVLCSGNGYLLGR
CCHSWKSGSECDVPTNQCIDTCSGHGTCTVGTICINPSYKGENCEEVDCLPTCSG
RGVCVGECHFCVGGPGCESPRASMEQCSGHGSFLADTNTCNDHNWTHDCSTE
LCAADCGHGICVAGSCDEBGMGTGCEQACHPRCSEHGCTCKDGKCECPGWGLD
CTTEGCLQNGNCTTGLNNGWYCVQCGWRGAGDTSMETACSGKNDGGLDTC
MDPDCCLQASCTTSLCVGSPDLIDIIQETQISSLSTQSFQYRHFELVGRDTHV
PDVNPFDGIHACVIRQVTSDDTGLVGNISINKPAYGTTIRQDGSFDLVNSGV
AIGLERAPPTQHEHMLPWRGFMDTIVMRHEVNDIPSCDLSFTPMPLVLA
PLTAPAGTCPERIGVPEIQTLOEEVRIIPGDMRLGYLSRTSGYKSLRLITLTHSTI
FESCDLILWEKVAEGLFRKFAAPNLSDYFQKVVGLTSRTSGYKSLRLITLTHSTI
PSPMLKMLHMAVAGLQVETASNLGWSVDKHALNIQSLHKGNGENFISQVPE
PVTGSLMNGRRSISPCSCNGADGNKLLAPALACGSDGSLYVGFVFRVFTIG
NVTSLVGLNSPAHKYLLATPVSWMYLSDTSSRKVKVKSILYAKVAKNLELVAG
TGQDCLPDETRCGDGGKAVEATLTPRGITVDKYGVIFFVDGTMTRIDQNGIISTL
LGNDRTPSARPLSCSDVMSIDQVRLEWPTDLAVPMDSLYLDNNVLIQISENHQVR
IVAGRPMHCQVGLDFELVSKIAIHATLESANALAVSHNGLLYIAESDEKINRVQV

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE

JOURNAL
TITLE

COMMENT
FEATURES
source

gene
CDS

STNGEISILAGAPSCDCNDANCDYSGDDGYAKDAKLNAPSSLAVSPOGELFIADL
GNIRIYVRNKAFLNPLNMEISSPIDDELYFDVNAASHVFTOSLTJTGGLYNFTYS
GEGDLSSITDKNKNRVISRRDSTGLPLMLMGPDGQTFWFTMGTNALKSAQAQGEIA
VMTYHSGSGLLATKSNEDGWSFYEYDNYGLTNTYTPGRVSSRTSDSSTVRQTE
GSNKEDITVTNLSASGFYTLQDQVKNYSYIIGLDSURLVLANGMEVSLHTEPHLL
SGTVNPTISKRNVTLPIDNGLNVEWRQKEQAVTVYGRRLRHNRLISMDFDR
VTRTEKYDDHRRKFTIRIHYDHAGFTLWAPSRLNGVNTVSPGGHIAIGIQTGMSV
RMEYDQNGRITSKIPADGKSWSYTYLEKSMVLLYSORQYIIEFDKNDLSSTWPNV
ARQTLTTRISGYRNTYRPPENATVLDQYSEDGLLQTIHQGTGRRVYIKGLSR
LLEILYDTRIAFSYDESAGMLKTVLQSEGFACIRIQIQLGLDRLQFRSEGMV
NAREPYNDSFRVTSQMAVNETPIPIIDYRDDVSGKTEQFGKGVYIYINQIT
TAVMTHTKHFDAYGRKVEQVEIFRSLMYMMVQDNMGVRVVAKELVGPYANTRYA
YEYDADGQLVVINDKPLMYRSGYLLNGLNLLSPNGARLPLADYDRLTRILGDV
QYRLDEDFLRQGRNDFEYNAGLLVTKYNGWNTIKYRVDGGRVRSSTOGHH
LQFFYADLSMPKTRVTHMYNHSSEITSPLYDLOGLHFAEMLSSGDEFVACDNITPL
AVFGAGIMIKQIILHTAFGEVLDNSPQLVIGYQGLYELPLKLVHMRDXYDLA
GRWTTDHDIDKRLNSDINVPNLYMFKNPNPLNSQETKMTIDNSWLVTFGPLY
NVIYGRKPTDAMEPSYELVHTQIKTOEMDSTKSVLQVQCEQRLKSPVRLERFGQ
IYSASDSGCPPTPLHTLTFATGSLFGKGVKVAIREGRVEADIISLANEDGRRIAAVLD
KASYLDQLHFTIAGLDTHYFVKGSLVEGDLILLGTMVGTORTLETCVNTVSVQNVVLG
KSRRIITDIOMQYGTLSLVRYSGSVDEKVRVLELQRAVATAWAHERHLRQGE
GSAWTDCGERQQLLSRGRVQYEGFYIVSDQFPELTDINNHNHFWQTEMGRR

ORIGIN

Alignment Scores:
Pred. No.: 2,156-50 Length: 9264
Score: 484.00 Matches: 87
Percent Similarity: 94.06% Conservative: 8
Best Local Similarity: 86.14% Mismatches: 6
Query Match: 88.16% Indels: 0
DB: 5 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AB026980 (1-9264)

Qy 1 ArgLeuPheArgTyrTrpPheAlaAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3806 AGACTTTTCGGAAGTGGTTTTCAGCTGCTCCCACTCTCTTATGACTTTGTTGGGAC 3865
Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3866 AAAACCCAGCTTACACCCAGAAAGTATATGGGCTGTCTGAGGCTTTTGTGTCGCGGG 3925
Qy 41 TyrGluTyrGluSerCysProAspLeuIleuTrpGluLysArgThrThrValLeuGln 60
Db 3926 TTTGAGTATGAATCCTCTCCAGATCTAATCTCTGGAGAAGCGGCGCTGTTCTGACG 3985
Qy 61 GlyTyrGluIleuAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 3986 GGATACGAGACCTGCGTCAAAATCTTTGGTGGCTGGAGTGTGGACAAACCATCGCGTG 4045
Qy 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 4046 AACATTCAAAGTGGTATCCTTTCACAGGCAACGGTGAGACATCTTCATTTCCACGACG 4105
Qy 101 Pro 101
Db 4106 CCT 4108

RESULT 7
AK122513 5804 bp mRNA linear ROD 15-MAR-2003
LOCUS
DEFINITION Mus musculus mRNA for mKIAA1455 protein.
ACCESSION AK122513
VERSION AK122513.1 GI:28972757
KEYWORDS FLI CDNA.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S.,
Nakajima,D., Nagase,T., Ohara,O. and Koga,H.
Prediction of the coding sequences of mouse homologues of KIAA
TITLE

Mon Aug 16 09:01:04 2004

us-10-029-020-14_copy_1100_1200.p2n.rge

gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries

DNA Res. 10, 35-48 (2003)

2 (bases 1 to 5804)

Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.

Direct Submission

Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research

Institute, Laboratory for Genome Informatics, 2-6-7

Kazusa-kamata, Kisarazu, Chiba 292-0818, Japan

(E-mail: mouse@kazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)

The CREATE program supported by Japan science and technology

corporation; cDNA full insert sequencing; Kazusa DNA Research

Institute; cDNA library construction, clone selection and 5'- &

3'-end one pass sequencing.

Location/Qualifiers

1..5804

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="mbg00055"

/tissue_type="brain"

/dev_stage="adult"

/note="vector:modified pBC SK"

1..5804

/gene="mKIAA1455"

<1..5488

/gene="mKIAA1455"

/note="CDS is predicted by in silico analysis. Start codon

is not identified."

/evidence="not experimental"

/product="mKIAA1455 protein"

/protein_id="BAC65795.1"

/db_xref="GI:28372758"

/translation="LIGVNVSLHSEYGYTTIRQDMFLVANGASLTLVPERSPF

LTCYHTVMTVNVVYMDTLVNMKENDIPSCDLGFRVPSPIVSPSTFRSSPE

DSPIIPETQLHEETIPGTDLKLSSRAAGYKSLKLTMTQAVIPNLMLVLMV

AVVGLQFQWFPASPLATYFIMDKTDANQKYGLSEAVSVGYEYESCLDTLWEK

RTVGLQYELDASMGWTLDRHVLVDQNGILYKNGENQFISQQPVPVSSVMNGR

RRSISCPNQAQGNKLLAPVALACIDGSLYRPFVYRFFPSGNVSVLELRNK

DRHSNPAHRYLATDPTVGLYVSDTNTIRYRPSKLTGAKDLTKNAEVAAGTEQ

CLFDFARCDGSKAVEATLMSPKMAIDKNGLIYFVDMGMRKQNGQIISLLGSN

DLTSAPLRCDSMHSIQLVLEMTDLAINPMDSIVLDNNVLOITENROVRIAG

RMHCQVGPVEYVPGHVAQTLLSATAINAVSISGLYITETDEKKNIRQVTTGE

ILSIVAGIPSECDKNDKADKINAPSSLAAPDGLIYADLGNRI

RAVKNKPLINSNFYVASTPQELIYFIDINGHQTVSLVTGLDLYNFSNDNDV

TAVIDNGNTLRIRRNPRVPSVDPNQVILITIGTNGSKMTAQGLELVFTYH

GNGLGATKSDTRGTFFDYDSEGRLTNTVFTGVNHLGDMDKAITVDIESRE

EDVSTISNLSIDSFYTMVDOLRNSYQIGYDGLSRIFYASGLDSHYQTPHVLGTA

NPTKARMLPDSGNQNLVEMFRKEQAKGVNFRKLRVNGRLNLSVDFDRTTKT

EKIYDHRKFLRLIADTSQHTLWLPSSKLMANVTYISGTQIASIQRTTSEKVDY

DSQGRIVSRVADKVTSTYLEKSMVLLHSQRYIFYEDMDRLSAITMPSVAHRT

MQTIRSGYRNTYNPNESNASITDNEEGLLQTLAFLTSRVLFPKRRQTRLSEI

YADLTPTRTYTHVNSSTISLYDLQGLFAMEISSGDEFYASDNTGTLPAVFS

SNGLMLKQIQTAYGSIYFSDNPQVDFVIGFHGGLYDPLTKLIHFGERDYDILAGWT

TPDLETKRIGKDPAPNLYMRNNPASKIHDVKYITDVSMLWTFFGHLNNAIPG

FVPKDELTPSVELKSOQMDVPPIFGVQOQVQAKAFSLGKMAEYQVSRKAG

AEQSWLWFAIVSLIGKGMALVSGRVQTNVLNANEDCIKVAALVNNAFYLENLHF

TIEGKDYFIKTTTPESDGLRLTSRKALENGINVTYSQSTVYVNGTRTFADVE

MQFGALHVRVYGMTLDEERKILQEARALAWAREQQRVDEEGEGARKLDEGEK

ROLLSAGKVOGYDGYVLSVEYQELADSNANNIQLRQSEIGKE"

ORIGIN

Alignment Scores:

Pred. No.: 7.98e-47 Length: 5804

Score: 454.00 Match: 79

Percent Similarity: 91.09% Conservative: 13

Best Local Similarity: 78.22% Mismatches: 9
Query Match: 82.70% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AK122513 (1-5804)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyPheIleTrpAsp 20

Db 494 AGACTCTTCCAGAGTGGTTCTCGCTCGCCAAACTTGGCCATACACGTTTCATCTGGGAT 553

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40

Db 554 AAGACGGACGCATATAATCAGAAAGTCACGGCTTGTACAGCGCAGTGTGTCTCCGTCGA 613

QY 41 TyrGluTyrGluSerCysProAspLeuLeuLeuTyrGluLysArgThrThrValLeuGln 60

Db 614 TACGAGTACGAGTCTGCTTGACCTGACTCTCTGGAAAGAGGACTCCGCTTTTGCAG 673

QY 61 GlyTyrGluLeuAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80

Db 674 GGCTATGAGTTCGATGCTTCGAACATGGCGCGCTGGACGTTGGACAAGCACCATTGACTG 733

QY 81 AsnIleGlnSerGlyTyrLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100

Db 734 GAGCTTCAGACGGTATCTACTATACAAAGGAATGGAGAAATCAGTTTCATCTCTCAGCAG 793

QY 101 Pro 101

Db 794 CTT 796

AF195418 7816 bp mRNA linear ROD 27-JAN-2000

Mus musculus ODZ3 (Odz3) mRNA, partial cds.

AF195418

AF195418.1 GI:6760368

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Ben-Zur, T., Feige, R., Motro, B. and Wides, R.

The mammalian Odz gene family: Homologs of a Drosophila pair rule

gene with expression implying distinct yet overlapping

developmental roles

Dev. Biol. 117, 107-120 (2000)

2 (bases 1 to 7816)

Ben-Zur, T., Motro, B. and Wides, R.

Direct Submission

Submitted (15-OCT-1999) Faculty of Life Sciences, Bar-Ilan

University, Ramat-Gan, Israel

Location/Qualifiers

1..7816

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

<1..7816

/gene="Odz3"

/note="synonym: Ten-m3"

<1..7042

/gene="Odz3"

/note="similar to the Drosophila melanogaster pair rule

odd Oz/Ten-m gene sequence"

/codon_start=2

/product="Odz3"

/protein_id="AAF28316.1"

/db_xref="GI:6760369"

/translation="NSTDTVNTVSLPSGDKGLGGFTNNNTTDSGLDGGRAIOE

VPGIFWRSQFLIDQPFQFLKFNLSQDALIGYVRKGLPSPHQYQFVELDGSRLI

AREQRNIVSERAGRSQARSLFFAGFIQYLDLSGIWHLAFYNDGNKRPQEVFNIVIE

SVVPCPRNGHNGSVGTCFPGFLGPDSCRAACVPVLCGNSGKACVCLCFSGMK

GTECDVTTCIDIPQCGRGICINGSCACNSGKYGENCEADCLDPCSNHGVCIHGE

Mon Aug 16 09:01:04 2004

us-10-029-020-14_copy_1100_1200.p2n.rge

MWDLRLSAITWPSVARHTMTIRSGYVNIYNPNESNAIIITDYNEGLLIQOTAFPLGT
 SRRVLYRROIRLSEILYDSTRVSFTTDETAGVLKTVNLQSDGFICTIRVQIGPLI
 DRQIFRSEDGMNARFDYSYDNRFRVTSQGVINETPLPDLIYQDIDSKRVQFQK
 FGVIYDINQIISTAVMTYKTFDAHGRKEIQEYIFRSLMYWITIYQDNNRVTKRE
 IKIGFANTTKVAYEYVDVQGLQVYLNKIMRWYNDLNGNLHLNPSARLITPLR
 YDLRRTIRLGDVOYRDEDEGLFQORGTETFEYSKGLLTRVYSKSGSWTVIYRYDGL
 GRRVSSITSLGHLQOFFVADLITYPRITHVYNHSSSETSLYLDQGLHFAEMLSSGD
 EYFJASDNTGTLAFSSNGMLKQITQTAIGEYIFDSDNDFQVLVPHGGLYDPLTX
 LIHFGEDYDILAGRTWTPDIEIKWRIGKDPAPENLFRNNPASKIHDVKDYITDV
 LSLWTFGHLHNAIPGPVPKFLDTPESYELVKSQWEDVPPIFGVQQVVARQAKFI
 LSKLWFGVHVRKAGAEOSLWLFATPKSLGKMLAVSQGRVQTVNLVIANEDKCI
 KVAVLNNAFLENLHFIEGKQDTHVFKITTPESDILGTLRLTSRKALENGINVTVS
 QSTTVNGTRFRFADVENQFQALAHVRYGTMILDEKARILEQORALARAWAREQO
 RVRDGEGARLWTEGKRLQSLSAGKVGQDGYVLVSVEQYPELADSANNIQFLRQSEI
 GKR"

ORIGIN

Alignment Scores: 1.27e-46 Length: 8964
 Pred. No.: 454.00 Matches: 79
 Score: 91.09% Conservativeness: 13
 Percent Similarity: 78.22% Mismatches: 9
 Best Local Similarity: 82.70% Indels: 0
 Query Match: 10 Gaps: 0
 DB:

US-10-029-020-14_COPY_1100_1200 (1-101) x AB025412 (1-8964)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheLeuTrpAsp 20
 Db 3203 AGACTCTTCAGAGTGGTTCTCGCTCGCCAACTTGGCTTACACGTTTCATCTGGAT 3262
 QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
 Db 3263 AAGACGACGACATATATCAGAAAGTCTACGGCTTCTCAGAGGACGTTGTCTCCGCGGA 3322
 QY 41 TyrGluTyrGluSerCysProAspLeuLeuLeuTyrGluLysArgThrThrValLeuGln 60
 Db 3323 TAGAGTACAGTGGTGGTGGACCTGACCTCTGGGAAAGAGGAGTGGCGGTTTGGCA 3382
 QY 61 GlyTyrGluLeuAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
 Db 3383 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3442
 QY 81 AsnIleGlnSerGlyLeuLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
 Db 3443 GACGTTTCAGAACGCTATCTATCAAGAAATGGAGAAATCAGTTTCATCTCTCAGCAG 3502
 QY 101 Pro 101
 Db 3503 CCT 3505

RESULT 10
 AK125869 3394 bp mRNA linear PRI 09-SEP-2003
 LOCUS Homo sapiens cDNA FLJ43881 fls, clone TEST14009028, highly similar
 DEFINITION to Mus musculus odd Oz/ten-m homolog 3 (Drosophila) (Odz3).
 ACCSSION AK125869.1 GI:34532128
 VERSION oligo capping; fis (full insert sequence).
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,
 Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
 Kawai-Hiro, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
 Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
 Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project

TITLE

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3394)
 AUTHORS Isogai, T. and Yamamoto, J.
 TITLE Direct Submission
 JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- and 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.

FEATURES

source

1..3394
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="TEST14009028"
 /tissue_type="testis"
 /clone_lib="TEST14"
 /note="cloning vector: pME18SFL3"

ORIGIN

Alignment Scores: 6.01e-47 Length: 3394
 Pred. No.: 453.00 Matches: 78
 Score: 91.09% Conservativeness: 14
 Percent Similarity: 77.23% Mismatches: 9
 Best Local Similarity: 82.51% Indels: 0
 Query Match: 9 Gaps: 0
 DB:

US-10-029-020-14_COPY_1100_1200 (1-101) x AK125869 (1-3394)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheLeuTrpAsp 20
 Db 1514 AGACTCTTCAGAGTGGTTCTCGCTCGCCAACTTGGCTTACACGTTTCATCTGGAT 1573
 QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
 Db 1574 AAGACGACGACATATATCAGAAAGTCTACGGCTTCTCAGAGGACGTTGTCTCCGCGGA 1633
 QY 41 TyrGluTyrGluSerCysProAspLeuLeuLeuTyrGluLysArgThrThrValLeuGln 60
 Db 1634 TAGAGTACAGTGGTGGTGGACCTGACCTCTGGGAAAGAGGAGTGGCGGTTTGGCA 1693
 QY 61 GlyTyrGluLeuAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
 Db 1694 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1753
 QY 81 AsnIleGlnSerGlyLeuLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
 Db 1754 GAGTACAGAACGCTATCTATCAAGAAATGGAGAAATCAGTTTCATCTCTCAGCAG 1813
 QY 101 Pro 101
 Db 1814 CCT 1816

RESULT 11

AB040888 5309 bp mRNA linear PRI 22-FEB-2001
 LOCUS Homo sapiens mRNA for KIAA1455 protein, partial cds.
 DEFINITION
 ACCESSION AB040888
 VERSION AB040888.1 GI:7959170
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Nagase, T., Kikuno, R., Ishikawa, K., Hirose, M. and Ohara, O.
 NEDO human cDNA sequencing project

Prediction of the coding sequences of unidentified human genes.
XVII. The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro
DNA Res. 7 (2), 143-150 (2000)
20277482
10819331
2 (bases 1 to 5309)
Chara, O., Nagase, T. and Kikuno, R.
Direct Submission
Submitted (04-APR-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL: <http://www.kazusa.or.jp/huge>, Tel: 81-438-52-3913,
Fax: 81-438-52-3914)

FEATURES

source
1. 5309
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fhl6070"
/clone_lib="pBluescriptII SK plus"
1. 5309
/gene="KIAA1455"
<1..>5309
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1455 protein"
/protein_id="BA049597.1"
/db_xref="GI:7959171"
/translation="QNOQPYCRGLPDPQDIISQSPQQAQKSFYDRIFLIGSDST
HVIPEGSPFNKSLASVIGVLTADGTPLIGVNVSPFHYEGYVITIQDGMEDLVAN
GGASLTIVFERSPELTQVHTVLPVNVVYMDTLVMEKENDIPSCDLSGVRNPIL
VSSPLSTFFERSPELPIETQVHEHTIPGDTLKLSTLSRAAGYKSVLKTMTQ
SIIPFNLMKVLHVAVVRGFLQKFPASPPLATYFIWDKTDVYNQKGLYSEAVSVG
YESCDLTLWEKRTAILQGYELDASNMGGWTLDKHVLVDQNGILYKNGENQFTS
QQPVSVMNGRRRSISCPSCNQADGNKLAPVALACIGDLSLVGDFNVVRRIF
PSGNVAVTELRNKRDFHNSNPARYVLTDPVTGDLVSDTNTTRLYRPKSLTGAKD
LTKNAEVSAGTGQCLPFBARCGGKAVEATLMSPKGNVAKNGLIYFVDTGMLRK
VDQNGIISTILGSDNLTSAELPCLDTSMTSHISQVLEWPTDLAINPMDNSIYLDNNVV
LOITENRQVTLAAGRPMHCVQVPEVPVGHVATQTTLESATAIAYSYSGLYITETDE
KKINRLQVTTGEISLAVGISECDKNDANCQDGYGADAKLSAPSLAASP
DGLYTLADLGNIRAVSKNPKLLNGMNFYVSPDQELYIFDINGHYTVSLVNG
DLYNFSYNDNDITAVTDSNGNTLRIIRDNPMPVRVSPDQVNLWLTGTLGKLS
MTAQGLELVLTFTYHSGILLATKSDDETTFDFDYSEGLTNTPTPTGVNTHGDM
DKAITVDIERSREEDVSIITSLSSIDSFTVTDQDLRSYQIGDGLKIIYASGLD
SHVQTEPHVLACTANTVAKRMWLPFGNQNLVWRFRKQQAQKSVFGRKLVNG
RLLSVDFRTTKETKIIDHRKFLRLIAYDTSHPITLMLPSSKLMVAVNVTYSTQI
ASIQRTTSEKVDYDQGRVSRVADGKTWSYILEKSMVLLHSQKVIFPYDMD
RLSAITMPSVARTMTQIRSIGYRNINYPESNASIIITDYNEEGLLQTAFLGTSR
VLFKYRQRLSEILYDSTRVSFTYDETAGVLKVNLDSDGFCITTYRQIGFLIDRQ
IFRFBGMVNAFDFYSDNSFVTSQGVINETPLIDLYQFDDISGKVGQFGRFV
IYVDINLIITAVMTYTKFDAGRIKEIQEYFRSLMTWITIQYDNMGRVTKREIKI
GFANITKAYEYVDQQLQTVLNEKIMRWYNDLGNLHNPNSARLTPLRYDL
RRIITKLGQVQLDEDELRQGTETIPEYSSKGLLTRYVSKSGWTVIYVDGLGR
VSSKTLGQHLQFFYADLYPTRIITHVNHSSSEITSLYDQLQHLFAMEISGDFY
IASDNTGLPLAVFSSNGLMKQIQTAYGEIYFVDSNIDQVLTGHEGGLYDPLTKLIH
FGERYDYLAGWTTPTDIEIKWIKDPAENLYMRRNNPASKHDDVKDITDVNSL
LVTEGFHLNLAIPGPVPREDITEPSYELVSKQDDIPPIFGVQVQQAQKAFSL
GKMAVQVSRRAAGQSLWFATVKSLLGKVMIAVSQGVTVQTVNLTANEDCIKVA
AVLNNAFYLENLHFTIEGDKTHFYIKTITTPESDLGLTLRLTSGR"

gene

1. 5309

CDS

1. 5309
/gene="KIAA1455"
<1..>5309
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1455 protein"
/protein_id="BA049597.1"
/db_xref="GI:7959171"

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
711 AGACTCTTCCAAAGTGGTTTCCTCCTCACCAACTGGCCTATACTTTTCATATGGGAT 770
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
771 AAACAGATGCATATAATCAGAAAGTCTATGCTCTATCTGAAGCTGTGTGTCAGTTGCA 830
QY 41 TyrGluTyrGluSerCysProAspLeuLeuLeuTyrGluLysArgThrThrValLeuGln 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
831 TATGATGATGAGTCGTGTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCAATTCGCAG 890
QY 61 GlyTyrGluLysAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
891 GGCTATGAATTGGATCGTCCAAACATCGGTGGCTAGCATAGATAAACAATCACCTGCTG 950
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
951 GATGTACAGAACCGGTATATCTGTACAAGGAAACGGGGAAACCACTTCTATCTCCAGCAG 1010
QY 101 Pro 101
|||||
1011 CCT 1013

RESULT 12

AX662357

LOCUS

Sequence 39 from Patent WO02062999.

DEFINITION

AX662357

ACCESSION

AX662357.1 GI:29163218

VERSION

1

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1

AUTHORS

Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,

Malyankar, U., Shinkets, R.A., Tchernev, V.T., Spaderna, S.K.,

Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,

Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,

Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,

Milliet, I. and Macdougall, J.R.

Proteins and nucleic acids encoding same

Patent: WO 02062999-A 39 15-AUG-2002;

Curagen Corporation (US)

FEATURES

Location/Qualifiers

1..8473

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1-6e-46

Score: 453.00

Length: 8473

Matches: 78

Percent Similarity: 91.09%

Best Local Similarity: 77.23%

Conservative: 14

Mismatches: 9

Query Match: 82.51%

Indels: 0

Gaps: 0

DB: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AX662357 (1-8473)

ORIGIN

1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20

3150 AGACTCTTCCAAAGTGGTTTCCTCCTCACCAACTGGCCTATACTTTTCATATGGGAT 3209

21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40

3210 AAACAGATGCATATAATCAGAAAGTCTATGCTCTATCTGAAGCTGTGTGTCAGTTGCA 3269

41 TyrGluTyrGluSerCysProAspLeuLeuLeuTyrGluLysArgThrThrValLeuGln 60

3270 TATGATGATGAGTCGTGTTGGACCTGCTGTGGGAAACGGGGAAACCACTTCTATCTCCAGCAG 3329

Mon Aug 16 09:01:04 2004

us-10-029-020-14_copy_1100_1200.p2n.rge

```

Qy 61 GlyTyrGluLeuAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db 3330 GCCTATGAATGGATGGCTCCAAACATGGTGGCTGGACATTAGATAAACATCATCGTGTG 3389

Qy 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3390 GATGTACAGAACGGTATCTACTGACAGGAACCGGGAAACCCAGTTTCATCTCCAGCAG 3449

Qy 101 Pro 101
Db 3450 CCT 3452

RESULT 13
LOCUS AX662359 8487 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 41 from Patent WO02062999.
ACCESSION AX662359
VERSION AX662359.1 GI:29163219
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
1 Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X.,
Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K.,
Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A.,
Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F.,
Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G.,
Millett,I. and Macdougall,J.R.
Proteins and nucleic acids encoding same
Patent: WO 02062999-A 41 15-AUG-2002;
Curagen Corporation (US)
FEATURES
source
1..8487
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1..6e-46 Length: 8487
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AX662359 (1-8487)

Qy 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3146 AGACTCTTCCAAAGTGGTTCTCGCTCACCACAACTTGGCTATATCTTTCATATGGAT 3205

Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3206 AAAACAGATGCATATATCAAGAAAGTCTATGGTCTATCTGAGCTGTGTGTCAGTTGA 3265

Qy 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGln 60
Db 3266 TATGAGTATGAGTCGTCTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATCTGCAG 3325

Qy 61 GlyTyrGluLeuAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db 3326 GCCTATGAATGGATGGCTCCAAACATGGTGGCTGGACATTAGATAAACATCATCGTGTG 3385

Qy 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3386 GATGTACAGAACGGTATCTACTGACAGGAACCGGGAAACCCAGTTTCATCTCCAGCAG 3445

Qy 101 Pro 101
Db 3450 CCT 3452

```

```

Db 3446 CCT 3448

RESULT 14
LOCUS AX662355 8645 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 37 from Patent WO02062999.
ACCESSION AX662355
VERSION AX662355.1 GI:29163217
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
1 Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X.,
Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K.,
Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A.,
Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F.,
Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G.,
Millett,I. and Macdougall,J.R.
Proteins and nucleic acids encoding same
Patent: WO 02062999-A 37 15-AUG-2002;
Curagen Corporation (US)
FEATURES
source
1..8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1..64e-46 Length: 8645
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AX662355 (1-8645)

Qy 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3322 AGACTCTTCCAAAGTGGTTCTCGCTCACCACAACTTGGCTATATCTTTCATATGGAT 3381

Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3382 AAAACAGATGCATATATCAAGAAAGTCTATGGTCTATCTGAGCTGTGTGTCAGTTGA 3441

Qy 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGln 60
Db 3442 TATGAGTATGAGTCGTCTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATCTGCAG 3501

Qy 61 GlyTyrGluLeuAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db 3502 GCCTATGAATGGATGGCTCCAAACATGGTGGCTGGACATTAGATAAACATCATCGTGTG 3561

Qy 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3562 GATGTACAGAACCGTATCTACTGACAGGAACCGGGAAACCCAGTTTCATCTCCAGCAG 3621

Qy 101 Pro 101
Db 3622 CCT 3624

RESULT 15
LOCUS AX662353 8675 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 35 from Patent WO02062999.
ACCESSION AX662353
VERSION AX662353.1 GI:29163216
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Spytek, K.A., Li, L., Wolenc, A.R., Vernat, C.A., Eisen, A., Liu, X.,
Malyankar, U., Shinkets, R.A., Tchernev, V.T., Spaderina, S.K.,
Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
Millet, I. and Macdougall, J.R.

TITLE
Proteins and nucleic acids encoding same

JOURNAL
Patent: WO 02062999-A 35 15-AUG-2002;

FEATURES

source
Location/Qualifiers

1..8675
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 1.64e-46 | Length: | 8675 |
| Score: | 453.00 | Matches: | 78 |
| Percent Similarity: | 91.09% | Conservative: | 14 |
| Best Local Similarity: | 77.23% | Mismatches: | 9 |
| Query Match: | 82.51% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-029-020-14_COPY_1100_1200 (1-101) x AX662353 (1-8675)

| | | | |
|----|------|--|------|
| QY | 1 | ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp | 20 |
| Db | 3334 | AGACTCTTCCAAAAGTGGTTTCTGCTCCTCACMAACTGGCTTACTTTTCATATGGGAT | 3393 |
| QY | 21 | LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly | 40 |
| Db | 3394 | AAAACAGATGCATATATCAGAAAGTCTATGCTGTAAGCTGTGTGTCAGTTGGA | 3453 |
| QY | 41 | TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGln | 60 |
| Db | 3454 | TATGAGTATGAGTCGTGTTTGGACCTGACTCTGTGGGAAAGAGGACTGCCATTTCTGCAG | 3513 |
| QY | 61 | GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu | 80 |
| Db | 3514 | GGCTATCAATTGGATGGCTCCACATGGTGGTGGACATTAGATAAACATCACGTGCTG | 3573 |
| QY | 81 | AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln | 100 |
| Db | 3574 | GATGTACAGACGGTATCTACTACAGGGAACGGGAAACCAAGTTTCATCTCCAGCAG | 3633 |
| QY | 101 | Pro 101 | |
| Db | 3634 | CCT 3636 | |

Search completed: August 14, 2004, 11:52:10
Job time : 1809.37 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 9568.56 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_1760_2300
Perfect score: 2845
Sequence: 1 YYIGADGSLRLILANGMEVA.....YNRAGSWSVRYDGLGRRV 541

Scoring table:

BLOSUM62
Ygapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029020@cgn 1 1 19065 @runat 06082004_112216_29275 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb.ba: *
2: gb.htg: *
3: gb.in: *
4: gb.ov: *
5: gb.pat: *
6: gb.ph: *
7: gb.pl: *
8: gb.pr: *
9: gb.ro: *
10: gb.ro: *
11: gb.ro: *
12: gb.sts: *
13: gb.sy: *
14: gb.un: *
15: gb.vi: *
16: gb.ba: *
17: em.fun: *
18: em.fun: *
19: em.in: *
20: em.mu: *
21: em.om: *
22: em.or: *
23: em.ov: *
24: em.pat: *
25: em.ph: *
26: em.pl: *
27: em.ro: *
28: em.un: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Query Length | DB ID | Description |
|------------|--------|---------|--------------|-------|--------------------|
| 1 | 2845 | 100.0 | 3048 | 9 | AK127705 Homo sapi |
| 2 | 2845 | 100.0 | 8354 | 6 | AX556500 Sequence |
| 3 | 2845 | 100.0 | 8624 | 9 | AB037723 Homo sapi |
| 4 | 2845 | 100.0 | 8645 | 6 | AX600210 Sequence |
| 5 | 2845 | 100.0 | 8993 | 9 | HSM806812 |
| 6 | 2840 | 99.8 | 3320 | 9 | AK127101 Homo sapi |
| 7 | 2834 | 99.6 | 3486 | 9 | HSM806114 |
| 8 | 2821.5 | 99.2 | 8438 | 6 | AX675551 |
| 9 | 2813 | 98.9 | 5583 | 10 | AK122490 |
| 10 | 2813 | 98.9 | 8585 | 10 | AB025413 Mus muscu |
| 11 | 2785 | 97.9 | 9722 | 10 | AF059485 |
| 12 | 2573 | 90.4 | 9264 | 5 | AB026980 |
| 13 | 2233 | 78.5 | 3614 | 6 | AX876525 |
| 14 | 2233 | 78.5 | 3614 | 6 | BD156175 |
| 15 | 2233 | 78.5 | 3614 | 9 | AK001336 |
| 16 | 2233 | 78.5 | 5309 | 9 | AB040888 |
| 17 | 2233 | 78.5 | 8473 | 6 | AX662357 |
| 18 | 2233 | 78.5 | 8487 | 6 | AX662359 |
| 19 | 2233 | 78.5 | 8645 | 6 | AX662355 |
| 20 | 2233 | 78.5 | 8675 | 6 | AX662353 |
| 21 | 2228 | 78.3 | 5804 | 10 | AK122513 |
| 22 | 2228 | 78.3 | 7816 | 10 | AF195418 |
| 23 | 2228 | 78.3 | 8964 | 10 | AB025412 |
| 24 | 2219 | 78.0 | 170156 | 2 | AP002412 |
| 25 | 2199.5 | 77.3 | 186084 | 9 | AP002768 |
| 26 | 2188 | 76.9 | 8816 | 5 | AB026979 |
| 27 | 2180.5 | 76.6 | 208760 | 2 | AC074048 |
| 28 | 2180.5 | 76.6 | 226256 | 10 | AC073599 |
| 29 | 2165 | 76.1 | 231970 | 2 | AC133064 |
| 30 | 2165 | 76.1 | 257693 | 2 | AC103323 |
| 31 | 2159 | 75.9 | 3270 | 6 | AX877449 |
| 32 | 2159 | 75.9 | 3270 | 6 | BD156663 |
| 33 | 2159 | 75.9 | 3270 | 9 | AK001748 |
| 34 | 2141.5 | 75.3 | 8797 | 6 | AX250063 |
| 35 | 2141.5 | 75.3 | 8797 | 6 | AX250066 |
| 36 | 2141.5 | 75.3 | 8797 | 6 | AX250066 |
| 37 | 2138.5 | 75.2 | 8409 | 5 | GA279031 |
| 38 | 2138.5 | 75.2 | 8409 | 6 | AX250068 |
| 39 | 2136.5 | 75.1 | 6560 | 6 | AX250065 |
| 40 | 2136.5 | 75.1 | 7781 | 9 | AB032953 |
| 41 | 2136.5 | 75.1 | 9729 | 6 | AX250013 |
| 42 | 2136.5 | 75.1 | 9826 | 6 | AX250008 |
| 43 | 2135.5 | 75.1 | 8689 | 6 | AX250067 |
| 44 | 2135.5 | 75.1 | 8689 | 10 | AF086607 |
| 45 | 2124.5 | 74.7 | 7713 | 9 | HSM808325 |

ALIGNMENTS

RESULT 1

AK127705 3048 bp mRNA linear PRI 09-SEP-2003
 LOCUS Homo sapiens cDNA FLJ45805 fis, clone NT2R13007095, highly similar
 DEFINITION to Mus musculus neurogranin 1 (Nrgr1).
 AK127705
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 AUTHORS
 Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S.,
 Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Watanabe, M.,
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
 Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
 Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A.,
 Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K.,
 and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 3048)
 REFERENCE
 AUTHORS
 Isogai, T. and Yamamoto, J.
 TITLE
 JOURNAL
 Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 252-0818, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 FEATURES
 source
 1. 3048
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2RI3007095"
 /cell_line="NT2"
 /cell_type="teratocarcinoma"
 /clone_lib="NT2RI3"
 /notes="cloning vector: pME18SFL3-mRNA from NT2 neuronal
 precursor cells treated 2-weeks mitotic inhibitor after
 5-weeks retinoic acid (RA) induction.-majorly NT2 neuron"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.01e-236 Length: 3048
 Score: 2845.00 Matches: 541
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-10-029-020-14_COPY_1760_2300 (1-541) x AK127705 (1-3048)
 QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValala 20
 Db 670 TACTACATCGGGCGCGATGGCTCTTGGGCTGTCTGGCCACACGGCATGGAGTGGCG 729
 QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
 Db 730 CTGCAGACTGAGCCCACTTCTGCTGGCTGGCCACCGCTCAACCCCGTGGGCAAGGAAT 789
 QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTtpArgGlnArgLysGluGln 60
 Db 790 GTACGCTGCCCATCGAACCGGCTCACTGTGGAGTGGCGCCAGCGCAAGAGCAG 849
 QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80

850 GCTCGGGCCAGGTCACTGTCTTTTGGGCGCGGCTTCGCGGTTCACAAACCGAAATCTCTTA 909
 QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
 Db 910 TCTCTGGACTTTGATCGCGTAAACGACAGAGAGATCTATGATGACCAACCGCAAGTTC 969
 QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrProSerSerArg 120
 Db 970 ACCCTTCGGATTCCTGACGACGCGGCGCGGCTCTGGTCAACCCAGCAGCAGG 1029
 QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
 Db 1030 CTGAATGGTGTCAACGTCACATCTCCCTCGGGGTTCATCTGCTGGCATCCAGAGGGGC 1089
 QY 141 IleMetSerGluArgMetGlnTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
 Db 1090 ATCATGTCTGAAGAAGTGAATACGACGAGGCGGCGGATCATCATCCAGGATCTTCGT 1149
 QY 161 AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
 Db 1150 GATGGGAAGACATGGAGCTACACATCTTAGAGAGTCCATGGTGTCTACTACACAGC 1209
 QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
 Db 1210 CAGAGCGAGTATATCTTTGAGTTCGACAAAGATGACCGCTCTCTTCTGTGACGATGCC 1269
 QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
 Db 1270 AACGTGGCGGCGAGACACTAGACCATCCGCTCAGTGGGTACTACAGAACTAT 1329
 QY 221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
 Db 1330 CAGCCCCCTGAGGGCAATGCCTCAGTCATACAGCACTTCACTGAGGATGGCACCTCTT 1389
 QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
 Db 1390 CACACCTTCTACCTGGGCACTGGCCGCGAGGTGATATACAAGTATGGCAACTGTCAAG 1449
 QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
 Db 1450 CTGCGAGAGACGCTCTATGACACCACCAAGGTTCAGTTTCCACTATGACGAGACGGCAG 1509
 QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
 Db 1510 ATGCTGAAGACCATCAACCTACAGAATGAGGGTTCACCTGCACCATCCGCTACCGTCAG 1569
 QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
 Db 1570 ATTTGGGCCCCCTGATTGACCGACAGATCTTCGGCTTTCACCTGAGGAGCATGGTCAAGCC 1629
 QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
 Db 1630 CGTTTTCATCAACATATGACACACGCTTCGCGGTGACCAAGCATGACGAGTGTGATCAAC 1689
 QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
 Db 1690 GAGACCCCACTGCGCCATTGATCTCTATCGCTATGATGATGTCAGGCAAGCATGGTCAAGCC 1749
 QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
 Db 1750 TTTGGGAAGTTTGGTGTCACTTACTATGACATTAACCGAGATCATCACCAGCTGTCTATG 1809
 QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
 Db 1810 ACCCACACCAAGCATTTTGTATGTCATATGGCAGGATGAAGAGTGCAGTATGAGATCTTC 1869
 QY 401 ArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
 Db 1870 CGCTCGCTCATGTACTCGATGACCGTCCAGTATGATATACATGGGCGGAGTAGTAGAAG 1929
 QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440

```

Db      1930 GAGCTGAAGTAGGACCCCTACGCCAATACCACTCGCTACTCTCTATGATGATGCTGAC 1989
QY      441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
Db      1990 GGCAGCTCGACACAGCTCTCCATCAATGACAGCCACTCTGGCGCTACAGCTACGACCTC 2049
QY      461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db      2050 AATGGGAACCTGCACCTTACTGAGCCCTGGACAGCTGCGGCTACACCACTACCGTAT 2109
QY      481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db      2110 GACATCCGCGACCGCATCATCTCGCTGGCTGACGTGCAATACAGATGATGAGATGCG 2169
QY      501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys 520
Db      2170 TTCCTGAGCAGCGGGCGGTGATCTTTGAGTACAACTCAGCTGGCGTGTCTCAAG 2229
QY      521 AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db      2230 GCCTACAAACCGGGCTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCGTGGCGGCGC 2289
QY      541 Val 541
Db      2290 GTG 2292

RESULT 2
LOCUS      AX556500      8354 bp      DNA      linear      PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION AX556500
VERSION    AX556500.1 GI:25899736
KEYWORDS   .
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS   Ganggoli, E.A., Patturajan, M., Vernet, C.A., Malyankar, U.M.,
            Kekuda, R., Stone, D.J., Anderson, D., Shinkets, R.A., Burgess, C.E.,
            Zerkhusen, B.D., Liu, X., Spyrek, K.A., Casman, S.J., Boldog, F.L.,
            Smithson, G., Li, L. and Ji, W.
TITLE      Polypeptides and nucleic acids encoding same
JOURNAL    Patent: WO 02057453-A 13 25-JUL-2002;
            Curagen Corporation (US)
FEATURES   source
            1..8354
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ALIGNMENT Scores:
Pred. No.:      3.81e-236      Length:      8354
Score:          2845.00      Matches:      541
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              6      Gaps:      0

US-10-029-020-14_COPY_1760_2300 (1-541) x AX556500 (1-8354)

QY      1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db      5312 TACTACATCGGGCGCGATGGCTCTCTCGGCTGCTGCTGGCCAAACGGATGGAGGTGGCG 5371
QY      21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db      5372 CTGCAGACTGAGCCCACTTGTCTGGCTGGCACCTCAACCCACCGTGGCGGACAGAGANT 5431
QY      41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60
Db      5432 GTCACGCTGCCCATCGCAACAGCGCTCAACCTGTGTGGAGTGGCGCCAGCGCAAGAGCAG 5491

```

```

QY      61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db      5492 GCTCGGGGCGAGTCACTGCTTTGGGCGCCGCTCGGGGTGCACACCGAAATCTCCTA 5551
QY      81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db      5552 TCTCTGGACTTTGATCGCGTAAACGACACAGAGAAGATCTATGATGACCACCGCAAGTTC 5611
QY      101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerSerArg 120
Db      5612 ACCCTTCGGATCTGTACGACCGCGGGGCGGCCAGCCCTCTGGTCAACCCAGCAGCAGG 5671
QY      121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db      5672 CTGAATGGTGTCAACGTCACATACCTCCCTCGGGGTTCATTTGCTGGCATCCAGAGGGGC 5731
QY      141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db      5732 ATCATGTCTGAAGAATGGAATACGACCGCGGGCGCCAGCCATCCATCCAGGATCTTCGT 5791
QY      161 AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db      5792 GATGGGAGACATGAGGCTACACATCTTAGAAGATCCATGGTGTCTGCTACTACACAGC 5851
QY      181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db      5852 CAGAGCGCAGTATATCTTTGAGTTTCGACAAGATGACCGCCCTCTCTTCTGTGACGATGCC 5911
QY      201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db      5912 AACGTGGCGCGGACACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTAT 5971
QY      221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db      5972 CAGCCCTCTGAGGGCAATGCCCTCAGTCATACAGGACTTCACTGAGATGGGCACCTCCT 6031
QY      241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db      6032 CACACCTTCTACCTGGGCACCTGGCGCGAGGTGTATATACAAAGTATGGCAAACTGTCAAAG 6091
QY      261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db      6092 CTGGCAGAGACGCTCTATGACACCAAGGTCAGTTTCACTATGACGAGACGCGCAGGC 6151
QY      281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db      6152 ATGCTGAAGACCATCAACCTACAGATGAGGGCTTCACTGCACCATCCGCTACCGTCAG 6211
QY      301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db      6212 ATTGGGCCCTGTGATTGACCGACAGATCTTCGGCTTCACTGAGGAAGGCATGTCGTCAG 6271
QY      321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db      6272 CGTTTTGACTACAACTATGACACAGCTTCGGGTGACCAAGCATGCGAGCTGTGATCAAC 6331
QY      341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db      6332 GAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTGAGGCAAGCAGAGCAG 6391
QY      361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db      6392 TTTGGGAAGTTTGGTGTCAATTACTATGACATTAACCCAGATCATCACACAGCTGTCTATG 6451
QY      381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db      6452 ACCACACCAAGCATTTTGTGATGATGATGAGGATGAAGAGTGTGAGTATGAGATCTTC 6511
QY      401 ArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db      6512 CGTCGCTCATGTACTGGATGACCGTCCAGTATGATTAACATGGGGCGAGTAGTGAAGAAG 6571

```


QY 201 AsnValalaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
 Db 1418 AACCTGGCGCGGAGACACTAGAGACCATCCGCTCACTGGCTACTACAGAAACATCTAT 1477
 QY 221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
 Db 1478 CAGCCCCCTGAGGCGCAATGCTCAGTCATACAGGACTTCACTGAGGATGGCACCTCTT 1537
 QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrIleTyrGlySerGlyLeuSerLys 260
 Db 1538 CACACCTTCTACTGGGCACTGGCCGAGGGTGATATACAAGTATGGCAAACTGTCAAG 1597
 QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
 Db 1598 CTGGCAGAGACGCTCTATCACACCAACCAAGGTCAGTTTCACCTATGACGAGCGGAGGC 1657
 QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
 Db 1658 ATGCTGAAGACCATCAACCTACAGAATGAGGGCTTCACCTGCACCATCCGCTACCGTCAG 1717
 QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
 Db 1718 ATGGGCCCTGATGGACGACAGATCTTCCGCTTCACCTGAGGAAGGCGATGTCAACGCC 1777
 QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
 Db 1778 CGTTTGTGACTACAATGACAAACAGCTTCGGGTGACCAAGCATGCAGGCTGTGATCAAC 1837
 QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
 Db 1838 GAGACCCCACTGCCATTTGATGTCATATGCTATGATGATGTGTGAGGAAGTGCATGAGATCTTC 2017
 QY 401 ArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
 Db 2018 CGCTCGCTCAGTACTGGATGACCGTCCAGTATGATAACATGGGCGAGTAGTGAAGAAG 2077
 QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
 Db 2078 GAGCTGAAGGTAGGACCTACGCCAATACCACTCGTACTCCTATGAGTATGATGCTGAC 2137
 QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
 Db 2138 GGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTC 2197
 QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
 Db 2198 AATGGGAACCTGCACCTTACTGAGGCCCTGGGAACAGTGCACGGCTCACACCACCTACGCTAT 2257
 QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
 Db 2258 GACATCCGCGACCGCATCACTCGCTGGGTGACGTGCATCAATCAAGATGGATGGATGGC 2317
 QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
 Db 2318 TTCTGAGGCGAGCGGGCGGTGATATCTTTGAGTACAACCTCAGCTGCGCTGCTCATCAAG 2377
 QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
 Db 2378 GCCTTACAACGGGCTGGCAGCTGGAGTGTGAGGTACCGCTACGATGGCTGGGCGCGC 2437
 QY 541 Val 541
 Db 2438 GTG 2440

RESULT 4

AX600210
 LOCUS AX600210 8645 bp DNA linear PAT 14-FEB-2003
 DEFINITION Sequence 22 from Patent WO2072830.
 ACCESSION AX600210
 VERSION AX600210.1 GI:28400252
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S.,
 Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,
 Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K.,
 Khare,R. and Wallia,N.K.
 TITILE Proteins associated with cell growth, differentiation, and death
 JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
 FEATURES
 source
 1..8645
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7489573CBI"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,98e-236 Length: 8645
 Score: 2845.00 Matches: 541
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-020-14_COPY_1760_2300 (1-541) x AX600210 (1-8645)
 QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
 Db 5361 TACTACATCGGGCGGCGATGGCTCTTGGGCTGTCTGTGGCCACCGCATCGAGGTGGCG 5420
 QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
 Db 5421 CTGCAGACTGAGCCCGCCCTTCTGGCTGGCACCGCTCAACCCACCGTGGCGCAAGAGGAAT 5480
 QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60
 Db 5481 GTCAAGCTGCCCATCGACACCGGCTCAACTGTGTGGAGTGGCGCCAGCGCAAGAGCAG 5540
 QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
 Db 5541 GCTCGGGGCGGAGGTCACTGTCTTTGGGCGCGGCTGGCGGTGGCAACACCGAAATCTCTCTA 5600
 QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
 Db 5601 TCTCTGACCTTTGATCCCGTACACGACAGAGAAGATCTATGATGACCCACCGCAAGTTC 5660
 QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrProSerSerArg 120
 Db 5661 ACCCTTCGGATTCTGTACGACACCGGGGGGCGCCAGCTCTGTGTCCACCGACGAGG 5720
 QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
 Db 5721 CTGAATGGTGTCAACGTGACATATCCCTTGGGGGTTCATATGCTGCATCCAGAGGGGC 5780
 QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
 Db 5781 ATCATGTCTGAAGAATATGAATACGACACGCGGCGCGCATCATCATCCAGATCTTCGCT 5840
 QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
 Db 5841 GATGGGAACACATGGAGCTACACATACTTAGAGAAGTCCATGGTGTCTGTCTACTACACAGC 5900
 QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200

| | | | | |
|----|------|-----|--|------|
| Db | 5901 | | CAGAGCGAGTATATCTTTGAGTTCCACAGAAATGACCGCTCTCTTCGTGACGATGCC | 5960 |
| QY | 201 | | AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr | 220 |
| Db | 5961 | | AACGTGGCGGCAGACACTAGACACCATCCGCTCAGTGGGCTACTACAGAAACATCTAT | 6020 |
| QY | 221 | | GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu | 240 |
| Db | 6021 | | CAGCCCCCTAGGGCAATGCCCTCAGTCATACAGGACTTCACCTGAGGATGGGACCTCCTT | 6080 |
| QY | 241 | | HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys | 260 |
| Db | 6081 | | CACACCTTCTACCTGGGCACTGGCGGAGGATATACAGATATGGCAACTGTCAAG | 6140 |
| QY | 261 | | LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly | 280 |
| Db | 6141 | | CTGGCAGAGACGCTCTATGACACCAAGGTCAGTTTCACCTATGACGAGACGGCAGGC | 6200 |
| QY | 281 | | MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln | 300 |
| Db | 6201 | | ATGCTGAAGACCATCAACTACAGATGAGGGCTTCACCTGCACCATCCGCTACCGTCAG | 6260 |
| QY | 301 | | IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla | 320 |
| Db | 6261 | | ATTGGGCCCCCTGATTACCGACAGATCTTCGGCTTCACTGAGGAGGCACTGTCAAGCC | 6320 |
| QY | 321 | | ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn | 340 |
| Db | 6321 | | CGTTTTGACTACACTATGACACAGCTTCGGGTGACGACGATGACGGGTGTGATCAAC | 6380 |
| QY | 341 | | GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln | 360 |
| Db | 6381 | | GAGACCCCACTGCCCATTCATCTCTATCGCTATGATGATGTGTCAGCAAGACAGACGAG | 6440 |
| QY | 361 | | PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet | 380 |
| Db | 6441 | | TTTGGGAAGTTTGGTGTCACTTATGATGATTAACAGATATCAACAGATCATCAGCTGATG | 6500 |
| QY | 381 | | ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe | 400 |
| Db | 6501 | | ACCACACCAAGCATTTTGATGATATGTCATATGTCAGGATGAAGATGTCAGTATGATCTTC | 6560 |
| QY | 401 | | ArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValValLysLys | 420 |
| Db | 6561 | | CGCTCGCTCATGTACTGATGACCGTCCAGTATGATTAACATGCGGCGGAGTAGTGAAGAAG | 6620 |
| QY | 421 | | GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp | 440 |
| Db | 6621 | | GAGCTGAAGGTAGGACCCCTACGCCAATACCATCTCGCTACTTCCATGATGATGATGCTGAC | 6680 |
| QY | 441 | | GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu | 460 |
| Db | 6681 | | GGCCAGCTGCAGACAGCTCCATCAATGACAAAGCACTCTCGCGCTACAGCTACGACCTC | 6740 |
| QY | 461 | | AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr | 480 |
| Db | 6741 | | AATGGGAACCTGCATCTACTAGCGCCCTGGGAACAGTGCACCGCTCACACCACTACGGTAT | 6800 |
| QY | 481 | | AspIleArgAspArgGlyThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly | 500 |
| Db | 6801 | | GACATCCGACCCGCACTCTCGCTGGTGCAGTGCAGTGCATCAAGATGATGATGATGCGC | 6860 |
| QY | 501 | | PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLysLys | 520 |
| Db | 6861 | | TTCTCTGAGCGACGCGGGCGGTGATCTTTGAGTACAACTCAGCTGGCGCTGCTCATCAAG | 6920 |
| QY | 521 | | AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg | 540 |
| Db | 6921 | | GCCTACAAACCGGCTGGGAGCTGGAGTGTGAGTACCGCTACGATCGCTGGGCGGCGC | 6980 |
| QY | 541 | Val | 541 | |

| | | | |
|------------|--|-------------|------|
| Db | 6981 | GTG | 6983 |
| RESULT | 5 | | |
| LOCUS | HS0806812 | | |
| DEFINITION | Homo sapiens mRNA; cDNA DKFP686K11107 (from clone DKFP686K11107). | | |
| ACCESSION | BX640737 | | |
| VERSION | BX640737.1 | GI:34364828 | |
| KEYWORDS | | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | 1 (bases 1 to 8993) | | |
| AUTHORS | Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S. | | |
| CONSRMT | The German Human cDNA Consortium | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY | | |
| COMMENT | Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFP686K11107) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/ . Location/Qualifiers 1..8993 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /map="llql4.1" /clone="DKFP686K11107" /tissue_type="human fetal kidney" /clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host DH10B, sites SfiIA + SfiIB" /dev_stage="fetal" 1..8993 /gene="DKFP686K11107" 1..4200 /gene="DKFP686K11107" /note="hypothetical protein, N-terminus elongated, differentially spliced" /codon_start=1 /product="hypothetical protein" /protein_id="CA545850.1" /db_xref="GI:34364829" /translation="RIDQNGIISTLLGSLNLTISARPLSCDSVMDISQVHLEWPTDLAI NPDNSLYLVLDNNVLIQISENHVRIIVAGHPMCOVPGIDHFLSLKVAIHATLESALA LAVSHNGVLIIAETDEKINIRIQVTTSGEISLVAGAPSCDCKNDANCCDFSGDDGY AKDAKLTPSLAVCAGELYVADLGNIRIRFIRKPKPFLNTQNMVSTGSLPDLWLPD FDTGKLYTQSLPTGTYTYNFTYTGDTITLTDNNNMVNVNRDSTGSLPDLWLPD GVVWVTMTGNSLTKSVYTGQHELAMTYHNSGLLTKNSASGAFYLLQDQVRSYII NVNPTQGVSSFRSDTDSVHVQVETSSKDDVTITNLSASGAFYLLQDQVRSYII CAGDSGLLLANGMEVALQTEPEHLLAGTVNPTVGRNVTLPIDNLGLNVEWRQKEQA RGQVTFGRRLRHVNRNLLSLDPDVTREKIVDHRKFTLRILLYDOAGRPSLWSPSS RLNGVNVTYSPGYIAGIQGINSERMEYDQAGRTSRIFADGKTWSTVYLKSMVLL LHSQRYIFTFKNDKRLSSVTMPNVAQTLTIRKSVGYIRNIYIYPPGNASVYIQDFT DGLLHTFYLGITGRVITYKGLSKLAETDYTTKVSFTYDETAGMLKTLINQEGFT CTYRQIGLIDQIRFETEGMVARFDYNDNSFRVTSQMAVINETPLIPDLIRY DDVSGKTEQKFGVYVDINQIITAVMTHTHFDAYGRMKQVETFRSLMYMTY QYDNKVRVKKELAVGPYANTRYSVEVDQGLQVTSINDPLWRVSYDNLGNLHLL SPGNSARVLPURIDIRIRLDGVQVQWEDGFLRQPGDI FEYNSAGLLIKAYNRA GWSVRYDGLGRVRSSKSHSLQFFYADLINPTKYTHLYNHSSEITSLYYDLQ GHLFAMELSSGDEFYIACDNTGFLAVFSGTGLMKQLIYTAIGLYMDTPNFQII YGHGGLYDPLTKLVHMGRRDYDLVLAGWTSFPHLWHLSSSNVNFNLYMFKNNPI SNSODIKCFMTDANSWLTTFQGLHNVIYPGPKPMDAMEPSYELIHTOMKTEWNDI KSILGVQCEVQKQLKAFVTLERFDQLYGSTTITSCQAQPKTKKFASSGSGVFGKGFAL | | |

KDGRVTTDIISVANEDRRVAAALINHAHYLENLHFTIDGVDTHFYVKPGPSEGLAIL
 GLSGRRITLNGVNVTVSOINTVNGRTRRYTDIOYCALCLNTRYGTTLDEERKARV
 LELARORARQAWARQQRLRGESEGLRATWTEGKKQVLSITGRVQYDGFVVISVEQY
 PELSDSANNIHFMWQSEMGRR"
 8953..8958
 /gene="DKFZp686K11107"
 8973
 /gene="DKFZp686K11107"

polyA_signal
 polyA_site

ORIGIN

Alignment Scores:
 Pred. No.: 4,19e-236 Length: 8993
 Score: 2845.00 Matches: 541
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x HSM806812 (1-8993)

QY 1 TyrTyrIleGlyValAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
 DB 1168 TACTATCATCGGGCCGATGGCTCTTTCGCGCTGCTGCTGGCCACCGCATGGAGTGGCG 1227
 QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
 DB 1228 CTGCAGACTGAGCCCCACTTGTGGCTGGCCGCTCAACCCACCGTGGGCAAGAGGAAT 1287
 QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
 DB 1288 GTCACGCTGCCCATCGACACCGCCCTCAACCTGTGTGGAGTGGCGCCAGCGCAAGAGCAG 1347
 QY 61 AlaArgGlyGlnValThrValPheGlyValArgLeuArgValHisAsnArgAsnLeu 80
 DB 1348 GCTCGGGCCAGGTCACTGTCTTTGGGCGCGGCTGCGGGTTTCAACCGGAATCTCCTA 1407
 QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
 DB 1408 TCTCTGGACTTTGATCGGTAAACAGCACAGAGAGATCTATGATCACCGCNAAGTTC 1467
 QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
 DB 1468 ACCCTTGGGATTCTGTACAGCAGCGGGCGGCCAGCCCTCTGTCTACCCAGCAGCAGG 1527
 QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
 DB 1528 CTGAATGGTGTCAACGTGACATATCTCCCTTGGGGTTTACATTGTGGCATCCAGAGGGGC 1587
 QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
 DB 1588 ATCATGTCTGAAGAATGGAATACGACAGCGGGCGGCATCATCCAGGATCTTCCT 1647
 QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
 DB 1648 GATGGGAAGACATGGAGCTACACATACTTAGAAGTCCATGGTGTCTACTACACAGC 1707
 QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro 200
 DB 1708 CAGAGGCAGTATATCTTTTGATTTCGAAGAATGACCGCTCTCTCTGTGACGATGCC 1767
 QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
 DB 1768 AAGGTGGCGGGCAGACACTAGACACCATCGCTCACTGGCTACTACAGAAACATCTAT 1827
 QY 221 GlnProGlnGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
 DB 1828 CAGCCCTTGGGGCAATGGCTCAGTCATACAGGACTTCACTGAGGATGGCACCTCCT 1887
 QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
 DB 1888 CACACCTTCTACCTGGGCACTGGCGCGAGGGTGATATACAAAGTATGCAAACTGTCAAAG 1947
 QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280

DB 1948 CTGGCAGAGACGCTCTATGACACCAAGGTCAAGTTCAGTTTACCTATGACGAGACGCGAGC 2007
 QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
 DB 2008 ATGCTGAAGACCATCAACCTACAGAATGAGGGCTTCACCTGCAACCATCCGCTACCGTCA 2067
 QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
 DB 2068 ATTGGGCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGCATGGTCAACGCC 2127
 QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
 DB 2128 CGTTTGTACTACAACTATGACCAACAGCTTCCGGGTGACCAAGCATGTCAGGCTGTGATCAAC 2187
 QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
 DB 2188 GAGACCCCACTGCCCATTTGATGCTATGATGATGTGTCTGAGCAAGACAGACGACAG 2247
 QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
 DB 2248 TTTGGGAAGTTTGGTGTCACTTACTATGACATTAACAGATCATCCACACGCTGTCTATG 2307
 QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
 DB 2308 ACCCACCAAGCAATTTGTATGTCATATGCAAGTGAAGAGTGCAGTATGAGATCTTC 2367
 QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValLysLys 420
 DB 2368 CGCTCGCTCATGTACTGGATGACCTCCAGTATGATAACATGGGGCGAGTAGTGAAGAAG 2427
 QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrTyrGluTyrAspAlaAsp 440
 DB 2428 GAGCTGAAGGTAGGACCTACGCAATACCACTCGTACTCTATGAGTATGATGCTGAC 2487
 QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
 DB 2488 GGCCAGCTGGCAGACAGTCTCCATCAATGACAAAGCACTCTGGCGCTACAGCTACGACCTC 2547
 QY 461 AsnGlyAsnLeuHisLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
 DB 2548 AATGGGAACCTTGCACTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACG 2607
 QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
 DB 2608 GACATCCGCGACCGCATCACTCGCTGGTGGTACGCTGCAATACAAAGATGATGAGATGGC 2667
 QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
 DB 2668 TTCTGAGGCGAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGCGCTGCTCATCAAG 2727
 QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
 DB 2728 GCCTACAAACCGGCTGCGAGTGTGAGGTACCGCTACGATGCGCTGGGGCGGCGC 2787
 QY 541 Val 541
 DB 2788 GTG 2790

RESULT 6

AK127101 3320 bp mRNA linear PRI 09-SEP-2003
 LOCUS Homo sapiens cDNA FLJ45158 fis, clone BRAWH304304, highly similar
 DEFINITION to Mus musculus neuregulin 1 (Nr1).
 AK127101
 ACCESSION
 VERSION AK127101.1 GI:34533862
 KEYWORDS Oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,

Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,
 Isono,Y., Kawai-Hio,Y., Sato,K., Nishikawa,T., Kimura,K.,
 Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
 Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
 Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K.,
 Masuko,Y., Nagai,K. and Isoqai,T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 3320)
 Isoqai,T. and Yamamoto,J.
 Direct Submission
 Submitted (15-JUL-2003) Takao Isoqai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.

FEATURES

Source

1..3320
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRAWH3043034"
 /tissue_type="brain"
 /clone_lib="BRAWH3"
 /note="cloning vector: pME18SFL3"

ORIGIN

Alignment Scores:

pred. No.: 3,06e-236 Length: 3320
 Score: 2840.00 Matches: 540
 Percent Similarity: 99.82% Conservative: 0
 Best Local Similarity: 99.82% Mismatches: 1
 Query Match: 99.82% Indels: 0
 DB: 9 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AK127101 (1-3320)

QY 1 TyrThrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
 DB 36 TACTACATCGGGCCGATGCTCTTGGCGCTGCTGCTGGCCACCGCATGGAGTGGCG 95
 QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
 DB 96 CTGCAGACTGAGCCCACTTGTGGCTGGCCACCGCTCAACCCCGTGGGCAAGGAAT 155
 QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGlnThrArgGlnArgLysGluGln 60
 DB 156 GTCACGCTGCCATCGACAACGGCTCAACCTGTGGAGTGGCCCGCCAGCAAGAGCAG 215
 QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
 DB 216 GCTCGGGCCAGGTCATGCTTTGGCGCCGCTGGCGGTTCACACCGAATCTCTTA 275
 QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleThrAspAspHisArgLysPhe 100
 DB 276 TCTCTGGACTTGTATCGGTAAACGACGACAGAGAAGATCTATGATGACCCGCAAGTTC 335
 QY 101 ThrLeuArgIleLeuThrAspGlnAlaGlyArgProSerLeuThrProSerSerArg 120
 DB 336 ACCCTTCGGATTCGTACGACGACGGCGGGCCCGCCAGCTCTGTGTACCCGACGACGAG 395
 QY 121 LeuAsnGlyValAsnValThrThrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
 DB 396 CTGAATGTTGTCAAGTGCACATCTCCCTCGGGGTTCATCTGCTGGCATCCAGAGGGGC 455
 QY 141 IleMetSerGluArgMetGluThrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160

DB 456 ATCATGTCTGAAAGAAATGGAATACGACACAGGGGGCCCATCATCATCCAGGATCTTCGCT 515
 QY 161 AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
 DB 516 GATGGGAAGACATGGAGCTACACATCTTACAGAGTCCATGGTGTCTGCTACTACACAGC 575
 QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
 DB 576 CAGAGGCAGTATATCTTTGAGTTCACAGAATGACCGCTCCCTTCTGTGTGACGATGCC 635
 QY 201 AsnValAlaArgGlnThrLeuGlnThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
 DB 636 AACGTGGGGCGGACACACTAGACCATCCCTCAGTGGGCTACTACAGAAACATCAT 695
 QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
 DB 696 CAGCCCCCTGAGGGCAATGCTCAGTCATACAGGACTTCACCTGAGGATGGCACCTCT 755
 QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
 DB 756 CACACCTTCTACCTGGGCACTGGCCGAGGGTGATATCAAGTATGGCAACTGTCAAAG 815
 QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
 DB 816 CTGGCAGACGCGCTCTATGACACCCACCAAGGTCACTTTCACCTATGACGAGCGGAG 875
 QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
 DB 876 ATGCTGAAGACCATCAACCTACAGATGAGGGCTTCACCTGACCATCCGCTACCGTCAG 935
 QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
 DB 936 ATTGGGCCCTGATTGACCGACAGATCTTCGCTTTCACCTGAGGAAGCATGGTCAAGCC 995
 QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
 DB 996 CGTTTTGACATCAACTATGACAAACAGCTTCGGGTGACCCAGCATGCGGCTGTGATCAAC 1055
 QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
 DB 1056 GAGACCCCATCGCCCATTTGATCTCTATGCTATGATGATGTGTGAGGAAGCATGGTCAAG 1115
 QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
 DB 1116 TTTGGGAAGTTTGGTGTCTATTTATGACATTAACAGATCATCATCCAGAGCTGTCTATG 1175
 QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
 DB 1176 ACCCACACCAAGCATTTTGTATGATGATGAGGATGAAGGAGTGCAGTATGATGATCTTC 1235
 QY 401 ArgSerLeuMetTyrThrMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
 DB 1236 CGCTCGCTCATGTACTGATGACCGTCCAGTATGATTAACATGGGGCGAGTAGTGAAGAAG 1295
 QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
 DB 1296 GAGCTGAAGTAGGACCCCTACGCCAATACCATCTCGCTACTCTTATGATGATGATGCTGAC 1355
 QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuThrArgTyrSerTyrAspLeu 460
 DB 1356 GCCCAGCTGCAGACAGTCTCCATCAATGACAAAGCCACTCTGGCGCTACAGTACGACCTC 1415
 QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
 DB 1416 AATGGGAACCTGCATCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCATCAGGTAT 1475
 QY 481 AspileArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
 DB 1476 GACATCCGCGACCGCATCTCCGCTGGGTGACGTGCAATACAAAGATGATGATGATGGC 1535
 QY 501 PheLeuArgGlnArgGlyGlyAspilePheGluTyrAsnSerAlaGlyLeuLeuLysLys 520
 DB 1536 TTCCTGAGCAGCGGGCGGTGATCTTTGAGTACAACTCAGCTGGCTGCTGCTCATCAAG 1595

```
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgGag 540
|||
Db 1596 GCCTACAAACCGGCGCTGGCAGCTGAGTGTGAGTACGCTACGCTGGCGCGCGCG 1655

QY 541 Val 541
|||
Db 1656 GTG 1658

RESULT 7
LOCUS HSM806114
DEFINITION Homo sapiens mRNA; cDNA DKFZp686D0412 (from clone DKFZp686D0412).
ACCESSION BX537983
VERSION BX537983.1 GI:31874053
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
Direct Submission
Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686D0412) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
1. 3486
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="11q14.1"
/clone="DKFZp686D0412"
/tissue_type="human cervix"
/clone_lib="f66 (synonym: hlcc3). Vector pSport1_sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
1. 3486
/gene="DKFZp686D0412"
<1. 3217
/gene="DKFZp686D0412"
/notes="hypothetical protein, N-terminus truncated"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAD97943.1"
/db_xref="GI:31874054"
/translations="SFGRLNVNTPPTQGVSSFRSDTSSVHVQVETSSKDDVTITNL
SASGAFVTLQDOVRNSYIYGADGSLRLLANGMEVALQTEPHLLAGTVNPTVGRKNV
TLPIDNLGLNVQRKEQAGQVTVFGRRLRVNHNLSLGFDRVTRTEKIYDDHRK
FTUKILYDQGRPSLWSPSSRLNGVNTVSPGGYIAGIQGIMSERMEYDQAGRTSR
IFADGKTWSTYILEKSMVLLHSQRYIFEFDKNDRLSSVTMPNVARQTLLETIRSVY
YRNIYOPPEGNASVIQDTEBDGHLHTFYLGTGRVRYLYKYGKLSKLAFTFYDTKYVF
TYDETAMLKTINLQNEGFTCTIRYQIGLIDRQIFRFTTEGWNARFDYNDNFR
VTSMAVINETPLPIDLYRXYDVSQTEQFGKGVIIYDINQIITAVMTHTKHFDAY
GRMKVEQYEI FRSLMTWMTVQIDNMGVRVKKELKVGFPYANTTIRSYEDADGLOTVS
INDKPLWRSYDINGLNLHLSPGNSARLPLRYDIRITRLGDVQKMDDEGLRQR
GGDIFPYNAGALLIKAYNRAGWSVRVRYDRLGRVSSKSHSHLQFFVADLTNPTK
VTHLYNHSSEISLSLYDQGHLPFAMELSSGDEFYIACDNIGTFLAVFSGTGLMIQOI
LYTAYGEIYMDTNPNFQIIGYHGGLYDPLTKLVHMRDLYDVLAVGRWSPDHELWKH
LSSSNVMPNLNFMKNNPILNSQDIKCFMIDVNSMLTFGFLQHNVIIPGPKPMDA
MTPSFLIHTOMKTQSDWNSKSLGVQCEVQKLFATLRFDLQYLSGTTISCOQAP
KEPFFASSGSVFGKGVFALKDGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTID
```

```
GVDTHTFVFKPSPSEGLAILGLSGGRTILENGVNVTVTSQINTVINGRTRYTIDIQY
GALCNTRYGTITLDEEKARVLELARQVRAQWARQORLREGEGLRAWTEGEKQOV
LSTGRVQGYDGFVFSVEQYPELSDSANNIHFMRQSEMR"
3468
/gene="DKFZp686D0412"

polya_site
ORIGIN
Alignment Scores:
Pred. No.: 1,08e-235 Length: 3486
Score: 2834.00 Matches: 539
Percent Similarity: 99.63% Conservatives: 0
Best Local Similarity: 99.63% Mismatches: 2
Query Match: 99.61% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_1760_2300 (1-541) x HSM806114 (1-3486)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
|||
Db 185 TACTACATCGGGGCGGATGGCTCTTGGCGCTGCTGCTGGCCACGGCATCGAGGTGGCG 244
|||
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAn 40
|||
Db 245 CTGCAGACTGAGCCCACTTGTCTGGCTGGCACCGTCAACCCACCGTGGGCAAGAGGAT 304
|||
QY 41 ValThrLeuProIleAspAsnGlyLeuValGluTyrArgGlnArgLysGluGln 60
|||
Db 305 GTACGCTGCCCATCGACACACGGCTCAACCTGTGGAGTGGCGCCAGCGCAAGAGGAG 364
|||
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeu 80
|||
Db 365 GCTCGGGGCGAGGTCACTGTCTTTGGCGCGCGCTGCGGGTTTCAACCGAAATCTCCTA 424
|||
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
|||
Db 425 TCTCTGGCTTTGATCGGTGAACACGACAGAGAGATCTATGATGACCCGCAAGTTC 484
|||
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerSerArg 120
|||
Db 485 ACCCTTCGGATTCTGTACGACGAGGCGGCGCCGCTCTGCTCACCCAGCAGCAGG 544
|||
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
|||
Db 545 CTGAATGGTGTCAACGTCACATCTCCCTGGGGGTTACATTGCTGGCATCCAGAGGGCG 604
|||
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
|||
Db 605 ATCATGTCGAAAGATGGAATACGACGCGGCGCGCATCATCCAGGATCTTCGCT 664
|||
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
|||
Db 665 GATGGGAAGACATGGAGCTACATACATCTAGAGAAGTCCATGGTGTGCTACTACACAGC 724
|||
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
|||
Db 725 CAGAGGCGATATATCTTTGAGTTCGACAAAGATGACCGCTCTCTCTGTGACGATGCC 784
|||
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
|||
Db 785 AACGTGGCGCGGACACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTAT 844
|||
QY 221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
|||
Db 845 CAGCCCTCTGAGGCAATGCTCAGTCATACAGGACTTTCATCTGAGGATGGGACCTCTT 904
|||
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
|||
Db 905 CACACCTTCTACCTGGGCACTGGCCGCGAGGTGATATACAAAGTATGCAAACTGTCAAAG 964
|||
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
|||
Db 965 CTGGCAGACAGCTTCTATGACACCACCAAGGTCAGTTTCCCTATGACGAGCGGAGGC 1024
```

[illegible]

| | | | |
|--|------|--|------|
| Db | 6038 | ACTGAGATGGGACCTCTTCACACCTTACCTGGCCACTGGCCGAGGTGATATAC | 6097 |
| QY | 254 | LysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPhe | 273 |
| Db | 6098 | AAGTATGGCAACTGTCAAGCTGGCAGAGACGCTCTATGACACCAAGGTCAGTTTC | 6157 |
| QY | 274 | ThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThr | 293 |
| Db | 6158 | ACCTATGACGAGACGGCAGCATGCTGAAGACCATCAACTACAGATGAGGGCTTACAC | 6217 |
| QY | 294 | CysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThr | 313 |
| Db | 6218 | TGCACATCCGCTACCGTCAGATGGGCCCTTGATGACCGACAGATCTCCGCTTCACT | 6277 |
| QY | 314 | GluGluGlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThr | 333 |
| Db | 6278 | GAGGAAGCATGGTCAACGCCGCTTTTGACTCACTATGACCAACAGCTTCCGGGTGACC | 6337 |
| QY | 334 | SerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAsp | 353 |
| Db | 6338 | AGCATGCAGGCTGTGATCAACGAGACCCCACTGCCATTCCTATCGCTATGATGAT | 6397 |
| QY | 354 | ValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGln | 373 |
| Db | 6398 | GTGTCAGGCAACAGACAGACGAGTTCGGGAAGTTTGGTGCTATTTACTATGACATTAACACG | 6457 |
| QY | 374 | IleIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLys | 393 |
| Db | 6458 | ATCATCACCAAGCTGTATGACCCACCAACCAAGCAATTTGATGATATGGCAGGATGAAG | 6517 |
| QY | 394 | GluValGlnTyrGluIlePheArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsn | 413 |
| Db | 6518 | GAAGTGCAGTATGATCTCCGCTCGCTCATGTACTGTGATGACCGCTCCAGTATGATTAAC | 6577 |
| QY | 414 | MetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyr | 433 |
| Db | 6578 | ATGGGGCGAGTAGTGAAGAGGAGCTGAAGGTAGGACCTACGCCAATACCACTCGCTAC | 6637 |
| QY | 434 | SerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeu | 453 |
| Db | 6638 | TCCTATGATGATGATGTCGACGGCAGCTGCAGACAGCTCCATCAATGACAAGCCACTC | 6697 |
| QY | 454 | TrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAla | 473 |
| Db | 6698 | TGGCGCTACAGCTACGACCTCAATGGGAACCTGACCTTACTGAGCCCTGGCAACAGTGCA | 6757 |
| QY | 474 | ArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGln | 493 |
| Db | 6758 | CGGCTCACACCTACCGTATGATCATCCGACCGCATCACTCGGCTGGGTGACGTGCAA | 6817 |
| QY | 494 | TyrlsMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsn | 513 |
| Db | 6818 | TACAAGATGATGAGGATGGCTTCTCAGGACGCGGGCGGTGATATCTTTGAGTACAAC | 6877 |
| QY | 514 | SerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArg | 533 |
| Db | 6878 | TCAGCTGGCTGTCTCATCAAGGCTCAACCGGCTGCGAGCTGGAGTGCAGTACCGC | 6937 |
| QY | 534 | TyrAspGlyLeuGlyArgVal 541 | |
| Db | 6938 | TACATGGCTTGGGCGGCGGTG 6961 | |
| RESULT 9 | | | |
| AK122490 | | | |
| LOCUS | | | |
| DEFINITION Mus musculus mRNA for mKIAA1302 protein. | | | |
| ACCESSION AK122490 | | | |
| VERSION AK122490.1 GI:28972711 | | | |
| KEYWORDS FLI, CDNA. | | | |
| SOURCE Mus musculus (house mouse) | | | |
| ORGANISM Mus musculus | | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1 Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S., Nakajima,D., Nagase,T., Ohara,O. and Koga,H.
Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries DNA Res. 10, 35-48 (2003)

2 (bases 1 to 5583)
Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
Direct Submission
Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatori, Kisarazu, Chiba 292-0818, Japan (E-mail:mousekazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing; Kazusa DNA Research Institute; cDNA library construction, clone selection and 5' - & 3'-end one pass sequencing.

Location/Qualifiers

1..5583

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="mbg04788"

/tissue_type="brain"

/dev_stage="adult"

/note="vector:modified pBC SK+"

1..5583

/gene="mKIAA1302"

<i..3597

/gene="mKIAA1302"

/note="CDS is predicted by in silico analysis. Start codon

is not identified."

/codon_start=1

/evidence=not_experimental

/product="mKIAA1302 protein"

/protein_id="BAC65772.1"

/db_xref="GI:28972712"

/translation="TONMYELSSPIDQLYLFDTSGKLYTOSLPTGDLNFTYTGDDITHIDNNKGMVNVPRDSTGMLVLPVDPGVYVMTGNSALRSVTTQGHLEAMM

TYGNSGLLATKSNENWTFEYDSEGRUINVTFTQGVSSPESDSDSSVHVQVETS

SKDDVTITNLASGAFYTLLOQVRNSYIGDGLSRLLELANGMEVALQTEPHLLAG

TVNPTGKRVNTPIDNLNLMVRQKQEQGVTFGRRLRVHNNRLSLDPDRVT

RTKXIYDHRKFTLRILYDQAGRPSPSSRLNGVNVTPSGHIAIGIORGIMSERM

EYDQAGTSIRIPADGKMWSTYVLEKSMVLHLSROVIFEFKNDRLSVWMPNVAR

OTLETIRSVGYRNIYOPPGNASVLODFTDGHLLHTFVLTGTRRVYIKYKLSKLA

ETLYDVTKVSFTYDETAGMLKTVNLQNEGTCTIRKQIGLIDRQIFRFTGGMVNA

RFYDYNDSFRVTSQMAVINETPLIDLYRYDDVSGKTEQFGKFGVYTDINQITTTA

VMTHIKFDAYGRMKVQVEIFRSLMYMTVQYDNMGRVVKELKVGFIYANTTRYSE

YDADGQLQTVSINDKPLMYSYDLNGLNHLSPNSARLTPRLYDLRDLTRLGDVOY

KMDDEGFLRQGGDFEYNSAGLLIKAYNRASGVSVRYVDGLGRRVSSKSSSHLQ

FFYADLTNPVTNTHYHNSSEITSLYDLOQHLFAMELSGGDFEYIACDNIGTFLAV

WTSPPHELKELSSNISIVFLHMYKNNPINSQDIKFMPTDVNSMLLTFQGLHNV

IPGTFPDDAMEPSYELVHTQMTQEWDNKSLIGVQCEVQQLKAFVLEFQDLRY

GSITTSQQAPETKPFASGSIKFGKQFKALDKGRVTTDIIISVANDGRRIALINNA

HYLENLHFTIDGVTFYKPGSEGLAILGLSGGRRTLENGVNVTVSINTMLSGR

TRYRTDIQYRALCLNTRYGTVDREKRVLELRQAVRQAWAREQQRLREGEL

RAWTDGKQQLVNTGRVQGVDFVTSVEQYPELSDSANNIHFMRQSGRR"

ORIGIN

Alignment Scores:

Pred. No.: 1..33e-233 Length: 5583

Score: 2813.00 Matches: 533

Percent Similarity: 99.26% Conservative: 4

Best Local Similarity: 98.52% Mismatches: 4

Query Match: 98.88% Indels: 0

DB: 10 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AK122490 (1-5583)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20

| | | | |
|--------------------------|------|--|-----------------|
| Db | 1645 | TTTGGGAAGTTTGGTGTCTATCTACGACATCAACCCAGATCATTTACCCAGCCGCTCATG | 1704 |
| Qy | 381 | ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluLeuPhe | 400 |
| Db | 1705 | ACCCACACCAAGCATTGTGATGCTTATGCGAGGATGAAGGAGTACAGTATGAGATTTTC | 1764 |
| Qy | 401 | ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys | 420 |
| Db | 1765 | CGGTCAAGAGTGGGACCTTCTGATGACTCTTCTAGTATGATAACATGGGACGGGTAGTGAAGAG | 1824 |
| Qy | 421 | GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp | 440 |
| Db | 1825 | GAGCTCAAGAGTGGGACCTTATGCCAACACTACCCGCTACTCTCTATGATGATGCTGAT | 1884 |
| Qy | 441 | GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu | 460 |
| Db | 1885 | GGCCAGCTGCACACAGTCTCCATCATGACAGCCACTCTGGGCTTACAGTATGACCTC | 1944 |
| Qy | 461 | AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr | 480 |
| Db | 1945 | AATGGGAACCTTACACTTGTGAGCCCTGGGAAACAGCGCTCACACCTACCGTAT | 2004 |
| Qy | 481 | AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly | 500 |
| Db | 2005 | GACCTCGTGACCGCATCTAGCTGGGTGGGTGATGATGATGATGATGATGATGATGATG | 2064 |
| Qy | 501 | PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys | 520 |
| Db | 2065 | TTCTTGAGGCGAGCGGGTGGGNTGCTTCGAGTACAATTCAGCGGGTGTCTCATCAA | 2124 |
| Qy | 521 | AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg | 540 |
| Db | 2125 | GCCTACAAACCGGGCTAGTGGGTGGAGTGTGAGTACCGATACGATGGTGGGACCGCA | 2184 |
| Qy | 541 | Val 541 | |
| Db | 2185 | GTA 2187 | |
| RESULT 10 | | | |
| AB025413 | | 8585 bp mRNA linear | ROD 08-MAY-1999 |
| LOCUS | | Mus musculus mRNA for Ten-m4, complete cds. | |
| DEFINITION | | AB025413 | |
| ACCESSION | | AB025413.1 | GI:4760781 |
| VERSION | | Ten-m4. | |
| KEYWORDS | | Mus musculus (house mouse) | |
| SOURCE | | Mus musculus | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| FEATURES | | | |
| Source | | | |
| 1. 8585 | | Location/Qualifiers | |
| /organism="Mus musculus" | | | |
| /mol_type="mRNA" | | | |
| /strain="Balb/c" | | | |
| /db_xref="taxon:10090" | | | |
| /tissue type="brain" | | | |
| /dev stage="adult" | | | |
| 1. 8585 | | | |
| /gene="ten-m4" | | | |
| gene | | | |

protein_id="AAC31807.1"
/db_xref="GI:3170615"
/translation="MDVKKRKYRSLTRRRDARRYTSSADSESGKPKQKSYSSST
LPHGYDQADLAYSRSRVKMPQAEFEFCRTGNTFTRELIGEMTTPHGTLYRTDIGL
KCGYSGMGSADLADLADVLPGLPEHVRLLWGRSTRSGSSCLSSRANSLTLTDEHE
NTGTGALHCSGASSTPIEQSPSPSPSPANESQRRLLHNGVAQPTPDSDEEFPN
SLVKGSGASLVAANDHSPSSQLNQHPRLATPPPLPHATPNQHHASINSLNQNGFT
PRNSPAPATLDSLSGEPFAGSAQETHAQMNVNLNSKTPVETRLNGKOPPLGTWQDN
LIEMDIFASRRDYGAYSDGHFFPKPGTSPLCITTSFGYVFMHLLFLNMLHMQEOMQ
FSPAPNLKPKSPYCNMKCAALISAILISATLILIAVFMHLLFLNMLHMQEOMQ
MYETIDETASSVPDVTGSLVPSGCTLETPDRKGAAGKSPSLFPDSSTIDSGEI
DVGRASQKTPCPTFWRSOVFIIDPHVHLKFNYSKGAALVIGYRGLKLPSPDLDV
ELIDGRLLTOEARSLEGQRQSRGPPSSHETGFIQYLDGSLWHILAFYNDGKGSV
VSLTIAESVNCNPNCSGDCISGTCFCELFELGPGDCGRASCPVLCGNGQYMKG
RCLCHSGMGAECVDPTNQCIDVACSSGTCIMGTICINPGYKGSCEBCEWDMCPCS
SRGVCVGHGVCVGGTCTCEDGMGAACQDACHPCRAEHGTCRDKCEKCEPHWNGE
EI CAADCGHGVCGGTCTCEDGMGAACQDACHPCRAEHGTCRDKCEKCEPHWNGE
HCTLCGPGCLNGNRCITLNLGNCVCLQGWRTGCDTSMETGCGDKNDGDLVD
CMFDCCQLPCHVCLCPDPLDIQTOAPVQOQLNFYDRIKFLVGRDSTHS
IPGNPDGHCACVIRGQVMSDGTPLVGNISFINNPLFGVTSRQDGSFOLVTVNG
ISILRPERFAPIQHEITLMLPMDRFVMEITVNRHEENELPSRDLSEARNPNVPS
SPLTSSASSCAEKGPVPEALQOESEIVTAGCKRRLYSLSRTPGKSVVRSLTHPT
IPNMLKHLVMAVEGRLPKFAAPDLSEYFIWDKTVYNQKFLVGSFAFVSQVQ
YESCPDLILWEKRTAVLQVEIDASKLGSMDKHLALNQSGILKHXGNGEVSQVQ
PVTGSLMGNGRBSISPCSNGLADGNKLLAPVALTCGSDGSLYGVDFNYTRIPPS
GNVTNLEHSHSPAKYLYLATDPMGSAFVLDSTNSRVEKVKSTTVVKDLVKNSEVA
GTQDCIPLFDTCRGGKATETLNPGRITVDKFLGLYFVDTGWTIRVDOQGLIST
LLGSDNLTARPVSCDSVMEISQVRLEWPTDLAINPMDSLYLIDNNVVLQISENHQV
RIVAGRPMHCQVPGIDQFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRQ
VTTSGEISLVAGAPSGCDKNDANCDFGDDGVAKADKLTSPSSLAVCADGSLYVAD
LGNRIREFIKNPFPLNARDSTGCMPLMLVPPDGVYVWTGNTNSALRSVTTQGHSL
AMTYSNGNSGLTAKSNGMTTFEYDYSFGLRINVTFTPGQVSSPSRSDTSVHVQV
ETSSKDDVTITLHSGSAGFYLLQDQNRNSYIIGADGSLRLHLLANGMEVALQTEPHL
LAVTNPVTKRNLPLTDINGLNLVEWRKEQAGQVTPGPRLRVHNNRLSLDLD
RVTEKTYDDHRRKTLRLYDQAGRSFMSPSRLNGVNTVYPCGHIIAGIQRMS
ERMEYDQAGRITSRIADGKMSYTYLEKMSVLHLHSQROVIFEDKNDRLSSWTMN
VARQLETIRSVGYRNIYQPPGNAVSIQDFTEDGHLHTFTYLGTRRVIYKYGLKS
KLAETLXTTTSVFTYDITAGKLVNLQNEGFTCTIRYQIGLIDRQVIFRTEGGM
VNAFNDYNSFRVTSQVAINETPILDIYDDYDDNGMVKELKLVGYANTYRI
TAVMTSHKHFDAYGRMEQVETFRSLMYMTVYDNDMGRVVKELKLVGYANTYRI
SVEYDADGQLOTVSINDKPLWRSYDGLNHLSPGNSARLTPLRDLRDRITRLGD
VQYKWDGQSLQRGDVFEVNSAGLLIKAVNRSAGSVRYRVDGLGRVSSKSSSH
HLQFYADLTNPTKTHLYSHSSSEITSLYDLOGLH.FAMELSGDEFFIACDNIGTP
LAVFSGDLMIXQILYTAIGYIYMDTPNFQIILGYHGLYDPLTKLIVHMGRDYDVL
AGRTSPDHEMLKRLSSNIVFFHLYMKNPNLSNQDIKCFMTDVNSWLLTGFQFL
HNVIPGYKPTDAMEPSYELVHTQMKTGQVNDNSKGLVQCEVKQKQKAFVTLERD
OLGYSTITSQCOAPEKTFASGSGIFGKGVFALKDGRVTTDIIISVANEDEGRRFAAIL
NNAHYLENHFTIDGVNTHYFVKPSPGDLAILLGSRRTELENGVNTVYQNTML
SGTRRYTDIQLQYRALCINRYGTVDEEKVRLVLELQARAVQAWAREQOREGE
EGLRAWTDGKQQLNTRGVQIDGFFVTSVEQVPELSDSANNIHFMRQSEMRGR"
1813.1881
/gene="Doc4"
/note="putative; transmembrane-region site"

misc_feature
7.38e-231 Length: 9722
Pred. No.: Matches: 528
Score: 2785.00
Percent Similarity: 98.52%
Best Local Similarity: 97.60%
Query Match: 97.89%
DB: Indels: 0
Gaps: 0

ORIGIN
Alignment Scores:
Pred. No.: Length: 9722
Score: 2785.00
Percent Similarity: 98.52%
Best Local Similarity: 97.60%
Query Match: 97.89%
DB: Indels: 0
Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AF059485 (1-9722)

Qy 1 TyrTyrileGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 6028 TACTACATCGGGCTGATGGCTCCCTGAGCTGCTACTAGCCAAATGCCATGGAAGTGGCT 6087
Qy 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyIysArgAsn 40
Db 6088 YTCAGACTGAGCCACACCTGCTGGCTGGCTGCTGCTCAACCCCACTGTAGCGAAGAAAT 6147

```

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluThrArgGlnArgLysGluGln 60
Db 6148 GTCACACTGCCCATTTAATAATGGCTCAACCTGGTGGAGTGGCGGAGCGCAAGGAGCAG 6207
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 6208 GCTCGTGGCGGAGCTCAGCTCTTTGACCCCGTTCGGGGTTTCAACACCGAAACCTCTTG 6267
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 6268 TCTTTTGACATTTGACCGTGTACACACACAGAGAGATCTTACCATGACCATCGCAAGTTC 6327
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuThrTrpSerProSerArg 120
Db 6328 ACCCTTCGATCCTATATGACACAGGAGGAGGCCAGCTTCTGGTCACTAGCAGCAGG 6387
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 6388 CTGAATGGTGTAAATGTGACCTACTCCCTGGAGGTCACATTGCTGGAAATCCAAAGGGGC 6447
QY 141 IleMetSerGluArgMetGluThrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 6448 ATCATGTCTGAGAGAAATGAATATGATCAGGAGGCGCATCACATCCCGGATCTTTGCA 6507
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 6508 GACGGGAAATGTGAGCTACAGTACTTAGAGAGTCCATGGTCTTCATCTCCACAGC 6567
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro 200
Db 6568 CAGAGGCAGTACATCTTCGAGTTTGACAAAGATGACCGCTCTCTCTGTGACCATGCC 6627
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 6628 AACGTAGCCGCGCAGACGCTGGAGACCATCCCTCAGTGGGTACTACAGGAACATCTAC 6687
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisIleLeu 240
Db 6688 CAGCCCCGGAGGCAACGCCCTCAGTCACTTCAGGACTTCACCTAGAGATGGACACCTGTA 6747
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 6748 CATACCTTCTACTTGGGCACCGCGCGGTGATTACAGTAGTGGCAGTGTGCAAG 6807
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 6808 CTGGCCGAGACTCTGTATGACACCATTAAGTAAAGCTTCACCTACGACGAGACCGCAGG 6867
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 6868 ATGCTGAAGACTGTCAACCTACAGATGAGGGCTTCACCTGCATATCCGTACCGTCAG 6927
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6928 ATTGGGCCCTGATTGATGGCAGATCTTCGGTTTACCGGAAGAGCATGTCATGCC 6987
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 6988 CGTTTTGATTACAACTATGACCAACAGTTTCGGTGTGACTAGCATGCGGCTGTATCAAT 7047
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db 7048 GAGACCCCACTGCCATTTGACCTTACCGCTATGATGATGTGTCAGGAAGACAGAGCAG 7107
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 7108 TTTGGGAAGTTTGGTGTCAATTACTACGACATCAACACAGATCATTACCACGCGCTCATG 7167
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 7168 ACCCACTCCCAAGCACTTTGATGTCTTATGGCAGGATGAAGGAAGTACAGTATGAGACTTTC 7227

```

```

QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 7228 CGGTCTACTCATGTACTGATGACTGTTCAGTATGATAACATGCGGCGGTAGTGAAGAAG 7287
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 7288 GAGCTGAAGGTGGAGCCCTATGCCAACACTACCCGCTACTCTCTATGAGTATGATGTAT 7347
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuThrArgTyrSerTyrAspLeu 460
Db 7348 GGCCAGCTGCAGACAGTCTCCATCATGACAGCCACTCTGGCGCTACAGCTATGACCTC 7407
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 7408 AATGGGAACCTACACTTGTGAGCCCTGGAAACAGCGCACCGCTCACACCATCTACGTTAT 7467
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 7468 GACCTCCGCTGACCGCATCACTAGGCTGGGTGATGTCTTCGAGTACAAATTCAGCCGGCTGCTCATCAA 7587
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLysLys 520
Db 7528 TCCCTGAGCGACGCGGTGGGATGTCTTCGAGTACAAATTCAGCCGGCTGCTCATCAA 7587
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 7588 GCCTCAATTCGGCTAGTGGGTGGAGTGTGAGTACCGATACGATGGGCTGGGACGCCGA 7647
QY 541 Val 541
Db 7648 GTA 7650

RESULT 12
AB026980
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. .9264
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/dev_stage="embryo"
1. .9264
/gene="ten-m4"
323. .8797
/gene="ten-m4"
/note="similar to Drosophila melanogaster tenn/odz and Mus musculus Doc4"
/codon_start=1

```

```

/product="ten-m4"
/protein_id="BAA81893.1"
/db_xref="GI:5307785"
/translation="MEVKRRPYRLSTRRTERRYSADSDGKINPKYSSET
LPHQDSRLVAGSVKDLVHEADFQRQDFSLRDMAFDPVPVPMHGYRLTDEML
KFAVSVASDADTDLGIMSPHVAFLMGRSNTKSGRSCLFSRANSNLTLTDEH
ENTENPPLHCSSASSPVDSYPPPSHAANOSQRLGNSGAQGRSDEDFGPN
SFLVTGSGNVCTPAAATANEGSFONHSLRTPPLPLPHSHSPSOHHTASIGLSRSN
YTRQNSPAPDTAPNEGPTSAQSSAQNWLNSVPLETRNIKAQTFLETLQD
NFEMDLATARDGAYTDGHLFRPGGTSPLYCTTSPGYPLTSVTSVPPRPLRN
TFSPAFSLKKYKHKCNMKCAALSALISVTLVFLAYFIAMHLFLANMLHPQVQRI
VOLTENDSLGLHLDLGLPLGNTGLEPDRGSDGKLDGFFDFSDIFDGEIDVG
RVAQLIPGIFWRQVFIHPMLKLVNLSKDALVGYGRGLPPSHPTDFDVELL
DERRLLSQGLDGLDPPFPAQQRSLVPIVTSHTDGLCYMDSGIMHLAVVNDCKETEYQ
SMFLTAISIDPCNPFNGDCVSGNCHCFPGFPGDCSRASCVPULSGNGQYLAGR
CMHSGWKGSECDVPTNQCIDITCSGHGTCIVGTICNPYSKGENEEDVCLDPTCSG
RGVCGEHCIFVWGPGCEPRASCEQCSGHGSLFADTNTCNCHNWTGHDCSTE
LCAADCGHGI CVAGSCRCDEGMGTGCEQACHPCRBHGTCKDKCSCPGWNGEH
CTIEGCPGLNGRGCTLNGWYCVCOLGWRGAGDTSMTACSDGDKNDGDGLTDC
MDPDCCLQSGCHTSLCVGSPDLIIQETQISSSLTQSFQRIHFLVGRDSTHVI
PDVNPDDGTHACVIRGQVVTSDGTPLVGNISFINKPAFYIITRODGSFDFLVSGGV
AIGLRPERAPFTIOTHTLWLPWRFFVMDTI VNRHEVNDIPSCDLSSTFRMPVLPLA
PUTAFAGTCPERIGVPEIQTLQEEVRI ECTDMRLGYLSRTSGYKSLIRLITHTSTI
PRLMKVHLMAVEGRFLFRKFTSAAPNLGYDFVMDTIVSYQKTVGLSFAFVSQGEY
ESCPDLILWEKTAVLQGYETTASNLGWSVDKHALNIQSLHKGNGENFIPISQOP
PVGTSMNGRRRSISCPSCNGLADGNKLLAPVALACGSDGSLVYGVDFNVVRI FTTG
NVTSLLENSPAHKYILATSPVSGWLYLSDTSRRKVKFKSLYAVKDAKNLELVAG
TGQCPLPYDTCRGGKAVEATLTPRGITVDKYGVIFFVDGTMIRRIDQGLISTL
LGFNDITSARPISCDVSMDI SQVLEWPTDLAVSPMDNSLYLDNNVLIQISNHQVR
IVAGRPMHQCVPGLDHFVSKIAIHAATLESANALVSHNGLIYIAESDEKKINRQV
STNGETISILAGASPDCNDKANDCVGDDGVAKDAKNAPSSLA VSDGELFIADL
GNIRIYVRNKAFLNPLNMYELSSPIDDELYLFEDVNAHVFTQSLTGTGDLNFYFS
GEGDLSITDKNKNRVSIRSDTGLPLMLPGQOTFPWMTGNTNALKSVAAGQOEIA
VMTYHSGSGLLATNEDGWSYFYDNYGRLNTVITYPTGRVSSYRTSDSIVRYOTE
GSKNEDITVNLASGTFYTLMDQDQXNSYYIGLDPGLSLVLANGMEVSLHTEPHLL
SGTVNPTISKRNVLPTDNLGNIWEVRQEQARQGVTVYGRRLVHNRLLSMDFDR
VTEKYVDHRRFTLRIHYDHAGRPTLWAPSRNLGNVTVSPGHHIAGIQRTMSV
RMEYDQNGRITSIKPADGKWSYTYLEKSMVLLYSORQVIFPEEDKNRLSVTWPV
ARQLETTSIGYRNTPEPNAQVLOYSDEGLLIQIHOGTGERVLYKYGKLSR
LLELYDTIRIASYDESAGMLKTVGLQSEGFACITRIYRQGLIDNRIQIRFRSEBGW
NARFDYNDSPFVTSQMAVINETPLIDLYRYDDVSGKTEQFGKFGVIYVDINQILT
TAVNTHKHFDVGRVKEVOYEIFRSLMYVMWQFDMGRVAVAKELKVGPIYANTRYA
YEYDAGOLQVRSNDKPLWYSYDLNGLNHLSPGNSARLTPLYRIDRIRTLGVD
QYRLDEGFLRQNDGDFEYNSAGLIVKTVNGVNTIKYRYDGLGRVSRSTQGH
LQFFYADLSPTVRTHMINASSSEITSLYDLOGLHFAELSSDSEFYVACDNIGTPL
AFVSGAGLMTKQILHTAFGEVYLDSPFQVIGVQGLYEPLTKLVHMGRRDYDLA
GRVTPDHDIRKILNSNIYFNLMEFNKNPNLSNETKCYMTDVNSMLVTFGFLY
NVTIPGVRKPTVDAMESYELVHTQIKTQEWDSKSVLQVQEVQRLKSFVRLERFQ
IYSASDSCPTPLHTLPTATGTSFGKGVKVAIRGRVEADIIISLANEDGRRIAAVL
KASYLDLHFTIAGLDTHYFVKGLVSGDLSILGTVGORTLETGVNVTVSQVNMVLG
GRSRRTIDIQVQGLTSLNRYGSSVDEEKVULEARQRAVATAMAHERRHLRQGE
GSRWTDGERQQLLSSGRVQGEYFIVSDVQFPFELIDTNINNVHFWQTEWRR"

```

ORIGIN

| | | | |
|------------------------|-----------|---------------|------|
| Alignment Scores: | | | |
| Pred. No.: | 1,65e-212 | Length: | 9264 |
| Score: | 2573.00 | Matches: | 475 |
| Percent Similarity: | 95.01% | Conservative: | 39 |
| Best Local Similarity: | 87.80% | Mismatches: | 27 |
| Query Match: | 90.44% | Indels: | 0 |
| DB: | 5 | Gaps: | 0 |

US-10-029-020-14_COPY_1760_2300 (1-541) x AB026980 (1-9264)

| | | |
|----|------|--|
| Qy | 1 | TyrTyrIleGlyAlaAspGlySerLeuArgIleuLeuLeuAlaAsnGlyMetGluValAla 20 |
| Db | 5765 | TACTACATCGACTGGATGGCTCTTTGGCTCTGCTGCTACTGGCAACCGCATGGATGCT 5824 |
| Qy | 21 | LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40 |
| Db | 5825 | CTCCACAGGAGCCCTCCTGCTGTCAGAACAGTGAACCCACCATAGCAAGCAAGAAC 5884 |
| Qy | 41 | ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60 |

| | | |
|----|------|---|
| Db | 5885 | GTACGCTGCCCATCGACATCGACTCAACCTGGTGGAGTGGAGACACGCGCAAGACGAG 5944 |
| Qy | 61 | AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80 |
| Db | 5945 | GGCGAGGCGCAGTCCAGTGTACCGACCGACACTCCGGGTTCATATAGAAATTTGTTA 6004 |
| Qy | 81 | SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100 |
| Db | 6005 | TCTATGGATTTCATCGGTAACACGAGTGTAGAAAGTGTATCATCATGACACACAGAGTTC 6064 |
| Qy | 101 | ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerArg 120 |
| Db | 6065 | ACCTTAAGCATCCATCATGCTCGAAGACCTTACATTTGGGCTCCAGTAGTCCCGG 6124 |
| Qy | 121 | LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140 |
| Db | 6125 | CTAAATGGTGTCAATGTCACTACTCACCGGTGGCCACATTCGTGTCATCCAGAGGGC 6184 |
| Qy | 141 | IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160 |
| Db | 6185 | ACAATGTCACTCGAATCGAATATGACCAAAATGGGAGAAATCACCTCAAAGATATTGCA 6244 |
| Qy | 161 | AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180 |
| Db | 6245 | GATGGAAATCATCGAGCTACACCTACCTTGAAGATGTCATGTGCTGCTCTACAGC 6304 |
| Qy | 181 | GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro 200 |
| Db | 6305 | CACGCAATACATCTTCGATTTGACAGAACGACCGCTCTTCTTCAGTAACATGCCC 6364 |
| Qy | 201 | AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220 |
| Db | 6365 | AATGTGCCACAGACAGACCTTGGAGACCCCGTTCCTTCCATGGCTATTACAGAAACACATAC 6424 |
| Qy | 221 | GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240 |
| Db | 6425 | CGGCCACCTGAGGCAATGCAAGTTTCCAGATTACAGTACAGTGAAGTGAAGTCTCTG 6484 |
| Qy | 241 | HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260 |
| Db | 6485 | CAAAACCATCCACCAAGAACAGGACGCTAGATCATCTATAAGTATGGAAATCTCCCGT 6544 |
| Qy | 261 | LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280 |
| Db | 6545 | CTGCTAGAAATCTCTTATGATACGACACGGAATTCCTTTCTTATGATGAATCTGCAGGC 6604 |
| Qy | 281 | MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300 |
| Db | 6605 | ATGCTCAAACTGTGGGCTTCAAGTGAAGTTTTCGTGTCATCATTACATCGTTACAGACAG 6664 |
| Qy | 301 | IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320 |
| Db | 6665 | ATCGACCACTTATCGACACACAGATCTTTCGTTTATGAGGAGGCGCATGGTAAATGCC 6724 |
| Qy | 321 | ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340 |
| Db | 6725 | CGCTTCGACTACAACTATGCAACAGCTTCGGGTCCACCATGCAAGCAGTAGTATTAT 6784 |
| Qy | 341 | GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360 |
| Db | 6785 | GAACACCCCTTGCCCATTTGCTATTCGCTATGATGACGCTCAGGCAAGCAGGAGCAG 6844 |
| Qy | 361 | PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380 |
| Db | 6845 | TTTGGCAATTTGGAGTTATATATTATGATATAACAGATATAACACACAGCTGTAAATG 6904 |
| Qy | 381 | ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400 |
| Db | 6905 | ACCACACCAACACTTTGATGTCATACGAAAGATTAAAGAGTGCAGTATGAGATTTC 6964 |
| Qy | 401 | ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420 |

Db 6965 CGTTCGCTCATGTATTGGATGGTACAGTTTGCACCAATGGGGCGAGTGGTGGCCAAG 7024

Qy 421 GluLeuIysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440

Db 7025 GAGCTTTAAAGTCGGACCCCTATGCAATACCAACCCCGCTATGCCTATGAATATGATGAGAT 7084

Qy 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrAspGlySerTyrAspLeu 460

Db 7085 GGTGAGCTCCAGTAGTGTCCATCAATGATAAACCCTCTATGGCGCTACAGTTATGACCTT 7144

Qy 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480

Db 7145 AATGGCAATTTCACCTCCTAGCCCTGATAGCCGCTGATAGCCGAGACTTACACCACTACGCTAC 7204

Qy 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500

Db 7205 GACATTCGGGACCGCATTAATCTCGTTCAGTGTATGTCAGTACCGTTTATGATGAGATGGA 7264

Qy 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520

Db 7265 TTCCTCAGCAAGAGGCAATGATTTCTTTGAGTATTAATCTGCTGGGTGCTTGTAAA 7324

Qy 521 AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540

Db 7325 ACCTACAAAGTTAACGGATGGACTATCAAGTACCGCTATGATGGCTTGGGCAGAGA 7384

Qy 541 Val 541

Db 7385 GTG 7387

RESULT 13

AX876525

LOCUS AX876525 3614 bp DNA linear PAT 17-DEC-2003

DEFINITION Sequence 11430 from Patent EP1074617.

ACCESSION AX876525

VERSION AX876525.1 GI:40031261

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

TITLE Primers for synthesising full-length cDNA and their use

JOURNAL Patent: EP 1074617-A 11430 07-FEB-2001;

FEATURES

source Location/Qualifiers

1..3614

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

149..3286

/note="unnamed protein product"

/codon_start=1

/protein_id="CAE89315.1"

/db_xref="GI:40031262"

/translation="MDKALIVDISSREEDSVITSNLSIDSPFTYVQDLNRSYOI

GYDGLRIITVAGSLDSHYQEPHVLAGTAMPTVAKRNMFLPGENQNLVEMFRKEQA

KLVNVRGRKLVNVRNLLSVDFDTTKTEKIYDHRKFLRLRAYDTSGHPTLWLPSS

KLVNVRGRKLVNVRNLLSVDFDTTKTEKIYDHRKFLRLRAYDTSGHPTLWLPSS

KLVNVRGRKLVNVRNLLSVDFDTTKTEKIYDHRKFLRLRAYDTSGHPTLWLPSS

LHSORVIFPEYDMMRLSALTMPVSARHTMTQIRISGYRNIYNPPESNAILTDYNE

EGLLQTAFTAGTDSRVLFKYRQTRLSLEILYDSRVSFTYDETAGVULKVNLQSDGFI

CTIRYQIGLIDRLQIFRSESDGMVNARFDYSIDNSFRVSMQGVNETPLPIDLYQF

DDLSGKVEQKGVIIYDINQIISTAVMTYTKHFDAGRIKEIQYEIFRSLMWITI

QYDNMGRTVETKIGFANTTKYAYEDVDQGLQTVLNEKIMRWYVNLNGLNHL

NPSNARLTPLRLDRIPLRGDVOVDELDFLORGTETPIEYSSKGLLTLYVSKG

SGMTVIRYDGLGRVRSKTSLSGHLQFFVADLTPTIRIYVNHASSETLSYLDIQ

GHLFAMEISSGDEFYASDNTGTPLAVFSNGLMLQIQIQTATGEIYFDSNIDFLVI

GFHGLYDPLTKLHGERDYDILAGRWITPDIEMKRIKGDPAFLNLYMFRNNPNS

KIHVDKDIYDVNSWLVTFGHLHNAIPGPPVFKDITPEYSELVKSQQWDDIPPIFG

VQOQVAKAFSLSGMAEVQVRRRAGGASQWLVATVKSILIGKVMVAVSQGRVQ

TNVLINANEDCIKVAALNNAFYLENLHFTIEGKDTHTYFIKTTTPESDLGTLTSGR

ORIGIN

Alignment Scores:

Pred. No.: 1.42e-183 Length: 3614

Score: 2233.00 Matches: 404

Percent Similarity: 88.91% Conservative: 77

Best Local Similarity: 74.68% Mismatches: 60

Query Match: 78.43% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AX876525 (1-3614)

Qy 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20

Db 272 TACCAGATTGGTTATGACGGCTCCCTCAGATATTATCTACGCCAGTGGCTTGGACTCACAC 331

Qy 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

Db 332 TACCACACAGAGCCGACGTTCTGGCTGGCACCGCTATCCGACGGTGGCCAAAAGAAC 391

Qy 41 ValThrLeuProIleAspAsnGlyLeuLeuValGluTyrArgGlnArgLysGluGln 60

Db 392 ATGACTTTTGGCTGGCGAGACGGTCAAAACCTTGGTGAATGGAGATTCCGAAAAGAGCAA 451

Qy 61 AlaArgGlyGlnValThrValPheGlyArgLeuArgValHisAsnArgAsnLeuLeu 80

Db 452 GCCNAGGGAAGTCAATGTTCTTTGGCGCAAGCTCAGGGTTAATGGCAGAAACCTCCT 511

Qy 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100

Db 512 TCAGTTGACTTTGATCGAACACAAAGACAGACAAAGATCTATGACGACCAACCGTAAATTT 571

Qy 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerSerArg 120

Db 572 CTACTAGGATCGGCTACGACACCGTGGGACCCGACCTCTCTGGCTGCCAAGCAGCAAG 631

Qy 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140

Db 632 CTGATGGCCGTCATGTCACCTATTCACAGGTCAAAATGGCAGCATCCAGCGAGGC 691

Qy 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160

Db 692 ACCACTAGCGAGAAAGTAGATTATGACGACAGGGAGGATCGTGTCTCGGGTCTTTGCT 751

Qy 161 AspGlyLysThrTyrSerTyrThrTyrLeuGlyLysSerMetValLeuLeuLeuHisSer 180

Db 752 GATGGTAAACATGAGAGTTACACATATTAGAAAAGTCCATGGTCTCTCTGCTTCATAGC 811

Qy 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200

Db 812 CAGCGCGAGTACATCTTCGATAGCATATGTTGGAGCCGCTGCTGCTCATCACCATGCC 871

Qy 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220

Db 872 AGTGTGCTCGCACACCATCGATCCGATCCATGGCTACTTACCCGCAACATATAC 931

Qy 221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

Db 932 AACCCCGGAAAGCAACCGCTCCATCATCAGGACTACACGAGGAGGAGGCTCTTCTA 991

Qy 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260

Db 992 CAACAGCTTCTTGGTACAGTCCGAGAGGCTTATTCAATACAGAGGAGAGTAGG 1051

Qy 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280

Db 1052 CTCTCAGAAATTTATATGATAGCACAAAGAGTCACTTACCTATGATGAAACAGCAGGA 1111

Qy 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300

Db 1112 GTCCCTAAGACAGTAACTCCAGAGTGGTGGTTTATTGTCACCATAGATACAGCAA 1171

Qy 301 lIeGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320

Db 1172 ATTGGTCCCTGATTGACAGGAGATTTCCGCTTTAGTGAAGATGGATGTAATGCA 1231

Qy 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340

Db 1232 AGATTGTTGACTATAGCTATGACACAGCTTTCCAGTGACAGCATGGAGGTGTGATCAAT 1291

Qy 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360

Db 1292 GAACGCCCACTGCTATGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1351

Qy 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380

Db 1352 TTTGGAAAGTTTGGAGTTAT 1411

Qy 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400

Db 1412 ACCTATACCAAGCACTTTGATGCTCAAGCCGCTATCAAGGAGATTCATATAGATATTC 1471

Qy 401 ArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValLysLys 420

Db 1472 AGGTGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1531

Qy 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440

Db 1532 GAGATTAAATAGGGCCCTTTGGCAACACCAACCAATATGCTTATGAATATGATGTGAT 1591

Qy 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460

Db 1592 GGCACACTCCAAAGCTTTACCTCAATGAAGAAATATGTTGGCGGTACACTACGATCTG 1651

Qy 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaAlaArgLeuThrProLeuArgTyr 480

Db 1652 AATGGAACTCCATTACTGAACCAAGTAACAGTGGCGGTCTGACACCCCTTCGCTAT 1711

Qy 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500

Db 1712 GACCTGCGACAGCAATCACTGACTGGGTGATGTTCAATATGCTTCCAGGGGCTTCACTCGA 1771

Qy 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys 520

Db 1772 TTCTACGTCAAAGGGGACGGAATCTTGAATATAGCTCCAGGGGCTTCACTCGA 1831

Qy 521 AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540

Db 1832 GTTTACAGTAAAGCGAGTGGCTGACAGTGTATCTACCGTTATGACGGCTGGGAAGCGT 1891

Qy 541 Val 541

Db 1892 GTT 1894

RESULT 14

BD156175

LOCUS 3614 bp DNA linear PAT 17-JAN-2003

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD156175

VERSION BD156175.1 GI:27861933

KEYWORDS JP 2002191363-A/11018.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K., and Otsuki, T.

TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL Patent: JP 2002191363-A 11018 09-JUL-2002;

COMMENT HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/11018

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORI

PI SAITO,

PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,

PI KEIICHI NAGAI, TETSUJI OTSUKI

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC

10, C12P21/02, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key

FT CDS Location/Qualifiers

(149)..(3283).

FEATURES

Source

1..3614

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1,42e-183 Length: 3614

Score: 2233.00 Matches: 404

Percent Similarity: 88.91% Conservative: 77

Best Local Similarity: 74.68% Mismatches: 60

Query Match: 78.49% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x BD156175 (1-3614)

Qy 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20

Db 272 TACCAGATTGGTTATGACGGCTCCCTCAGAAATTTATCTACGCCAGTGGCTGGACTCACAC 331

Qy 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

Db 332 TACCAACAGAGCGGCACCGTTCTGGCTGGCCCGCTAATCCGACGGTTCGCCAAAGAAC 391

Qy 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60

Db 392 ATGACTTTGCTGGCGAGAACCGTCAAACTTGGTGGATGGAGATTCGCAAAAGAAC 451

Qy 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80

Db 452 GCCCAAGGGAAGTCAATGTTCTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCT 511

Qy 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100

Db 512 TCAGTTGACTTTGATCGAACACACAGACAGAAAGATCTATGACGACACCGTAATTT 571

Qy 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerArg 120

Db 572 CTACTGAGGATCGCTACGACACGCTCTGGCCACCGACTCTCTGGCTGCCAAGCAGCAAG 631

Qy 121 LeuAsnGlyValAsnValThrTyrSerProGlyTyrIleAlaGlyIleGlnArgGly 140

Db 632 CTGATGCCCTCAATGTCACTTATTCACAGGTCAAAATGGCCATCCACGACGAGGC 691

Qy 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160

Db 692 ACCACTAGCGAGAAAGTAGATTATGACGACGAGGAGGATCGGTCTCGGCTCTTTGCT 751

Qy 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180

Db 752 GATGGTAAACATGAGTTCACATATTTAGAAAAGTCCATGGTCTCTTCTGCTTCATAGC 811

Qy 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro 200

Db 812 CAGCGGAGTACATCTTCGAATAGATATGACGACGAGGAGGATCGGTCTCGGCTCTTTGCT 871

Qy 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220

Db 872 AGTGTGGCTGCCACACACCATCGACAGCATCCGATTCCTGCTACTACCGCAACATATAC 931

QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
 Db 932 AACCCCGGGAAGAACACCCCTCCATCATCAGACTACAGGAGGAGGCTGCTTCTA 991
 QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
 Db 992 CAACACAGCTTCTTGGGTCAAGTCGAGGGCTTATTCAATACAGAGGAGGAGCTAGG 1051
 QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
 Db 1052 CTCTCAGAAATTTATATGATGACACAGAGTCAGTCTTTACCTATGATCAACAGCAGCA 1111
 QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
 Db 1112 GTCTTAAGACAGTAACCTCCAGAGTCAGTGTATTATTTGCACCATTAGATACAGGCA 1171
 QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
 Db 1172 ATTTGTCCTCCCTGATTCACAGGCAGATTTTCCGCTTTAGTGAAGATGGGATGTAATGCA 1231
 QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
 Db 1232 AGATTGTACTATGACTATGACACAGCTTTCAGTGCACGAGCATGCAGGTCGATCAAT 1291
 QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
 Db 1292 GAAAGCCACTGCTATGATCTGTATGATCTGATGATGATGATGATGATGATGATGAT 1351
 QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrAlaValMet 380
 Db 1352 TTTGGAAGTTTGGAGTTATATATATGATATTAACACAGATCAITTTACAGCTGTAATG 1411
 QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
 Db 1412 ACCATATGAAAGCACTTTGATGCTATGATGCTATGATGATGATGATGATGATGATGAT 1471
 QY 401 ArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
 Db 1472 AGGTGCTCATGTACTGATTAACATTCAGTATGATTAACATGCGGTGCGGTAAACAGAGA 1531
 QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
 Db 1532 GAGATTAATAATAGGCGCCCTTTGCCAACACACACCAATATGCTTATGATATGATGATGAT 1591
 QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
 Db 1592 GGACAGCTCCAAACAGTTTACCTCAATGAAAGATAATGTGGCGGTACACATCAGATCTG 1651
 QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
 Db 1652 AATGGAAACCTCCATTTACTGAACCAAGTAACAGTCGCGGTCTGACACCCCTTCGCTAT 1711
 QY 481 AspileArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
 Db 1712 GACCTCGACACAGAAATCACTCGACTGGGTGATCTTCAATATCGTTGGATGAAGATGGT 1771
 QY 501 PheLeuArgGlnArgGlyGlyAspilePheGluTyrAsnSerAlaGlyLeuLeuLysLys 520
 Db 1772 TTCTTACGTCAAAGGGGACCGGAATCTTTGAATATAGCTCCAAAGGGGCTTCAACTCGA 1831
 QY 521 AlaTyrAsnAlaGlySerTyrPheValArgTyrArgTyrAspGlyLeuGlyArgArg 540
 Db 1832 GTTTACAGTAAGCAGCAGCTGCTGACAGTATCTACCGTTATGACGCCCTGGGAAGCGGT 1891
 QY 541 val 541
 Db 1892 GTT 1894

RESULT 15
 AK001336
 LOCUS Homo sapiens cDNA FLJ10474 fis, clone NT2RP2000067.
 DEFINITION

AK001336
 VERSION GI:7022530
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosoiri,T., Raku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
 NEDO human cDNA sequencing project
 TITLE Unpublished
 JOURNAL 2. (bases 1 to 3614)
 REFERENCE Isogai,T. and Otsuki,T.
 AUTHORS Direct Submission
 TITLE Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 JOURNAL NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
 FEATURES
 source Location/Qualifiers
 1. 3614
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2RP2000067"
 /cell_line="NT2"
 /cell_type="teratocarcinoma"
 /clone_lib="NT2RP2"
 /note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."
 149..3286
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAA91633.1"
 /db_xref="GI:7022531"
 /translation="MDKAITVDIESSRDVSITSNLSIDSYTVMQDLNRSYQI
 GYDGLRIIYASGDSHYQTEPHVLAGTANPTVAKRMTLPGENQNLVWRFRKEQA
 QGKVNPGRLVNGRNLNSVDFDTTKTEKIYDHRKFLRLIAYDTSGHPTLWPS
 KLVANVTYSSTGTQIASIQRTTSKVDYDQGRIVSRVADGKTWSTYVLEKSMVLL
 LHSQRYIFEDMDRLSALTTPSVARHTMTGIRSYENIYNPPESNASTIDVNE
 EGLLQTLAFLGTSRRVLKRYRQTLSEILLYDSTRVSFTYDETAGVLKTNVLAGDGI
 CTRRYQIGPLLDRLQIFRSEDGNVARFDYSDNSFRVTSMQGVINETPLIDLYOF
 DDISKVEQKFGVIYIDNQIISTAVMTYTKHFDAGRIKEIQEIEFRSLMYWITI
 QDNMRVTKREIKGFANFTKAYEYVDVQGLQTVLYLNEKIMRWYNYDNLGNLHL
 QPNSARLPLRYDLDRITRLDQVDFRLEDGFLQRGTETIPEYSSKGLTLTVSKG
 SGMVTIYRVDGLGRVRSKSTLQHLQFFVADLTPTRLTHVYNHSSSETLSLYDLQ
 GHLFAMEISGDEFFYASDNTGTPPLAVFSSNGLMLKQIQYTAIGEYIFDNIIDFLVI
 GEHGLYDPLTKLHLFGERDYDLAGRTWTPDIEIKWRIKQIPAPNLNFRNNFAS
 KHDVADYITDVNSLWLTFFGLHNAIPGFPVPKFDLTPEFSLVKSQQMDDIPPIFG
 VQOVARQAKAFSLGKMAEYQVRRRAGCAQSLWFAVTKSLIGKVMVAVSQGRVO
 TNLINIANEDCIKAAVLNNAFYLENLHFTIEGKTHYFKTKTPSDGLTLTSGR
 KALENGINTVTSQSTTVNNGRTFRPADVEMOFGALAHVRYGTMLEEKARILEQARQ
 RALARAREQQRVDRGEGSARLWTEGKRLLSACKVQGDYGYVLSVEQYPELADS
 ANNIQPLRQSEIGRR"

CDS

ORIGIN

Alignment Scores:
 Pred. No.: 1,42e-183 Length: 3614
 Score: 2233.00 Matches: 404
 Percent Similarity: 88.91% Conservative: 77
 Best Local Similarity: 74.68% Mismatches: 60

Query Match: 78.49% Indels: 0
 DB: 9 Gaps: 0
 US-10-029-020-14_copy_1760_2300 (1-541) x AK001336 (1-3614)
 QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
 DB 272 TACCAGATTGGTATGAGCGCTCCCTCAGAAATTTCTACGCCAGTGGCTGGACTCACAC 331
 QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
 DB 332 TACCAACACAGACCGCACGTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAGAAAC 391
 QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluThrArgGlnArgLysGluGln 60
 DB 392 ATGACTTTTGGCTGGCGAGAACCGTCAAACTTGGTGGAAAGGAGATTCGAAAGAGCAA 451
 QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuValHisAsnArgAsnLeuLeu 80
 DB 452 GCCCAAGGGAAGTCAATGCTTTTGGCCGACGCTCAGGGTTAATGGCAGAAACCTCCT 511
 QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAsnHisArgLysPhe 100
 DB 512 TCAGTTGACTTTGATCGAACAACAAGACAGAAAGATCTATGACCAACCCGCTAAATTT 571
 QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuThrSerProSerArg 120
 DB 572 CTACTGAGGATCGCTTACGACAGCTCTGGGCAACCGACTCTCTGGTCCCAAGCAAG 631
 QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
 DB 632 CTGATGGCGTCAATGTACCTTATTCATCCACAGGTCAAAATGCCAGCATCCAGCGAGGC 691
 QY 141 IleMetSerGluArgMetGluThrValAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
 DB 692 ACCACTAGCGAGAAAGTAGATTATGACGGACAGGGAGGATCGTGTCTCGGGTCTTTGCT 751
 QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
 DB 752 GATGGTAAACATGGAGTTACACATATTTAGAAAGTCCATGGTTCTTCTGCTTTCATAGC 811
 QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro 200
 DB 812 CAGCGGAGTACATCTTCGATACGATATGTTGGACCGCTGTCTGCCATCACCATGCC 871
 QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
 DB 872 AGTGTGGCTCGCCACACCATCGACACCATCCGATCCATGGCTACTACCGCAACATATAC 931
 QY 221 GlnProGluGlyAsnAlaSerValIleGluAspPheThrGluAspGlyHisLeuLeu 240
 DB 932 AACCCCGGAAAGCAACGCTCCATCATCGGACTACACGAGGAGGGCTCTTCTTA 991
 QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
 DB 992 CAACACGCTTCTTGGGTACAGTCGAGGGTCTTATTCAATACAGAAAGGCACTAGG 1051
 QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
 DB 1052 CTCTCAGAAATTTTATATGATAGACAGAGTCTAGTTTACCTATGATGAACAGCAGCA 1111
 QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
 DB 1112 GTCTTAAAGACAGTAACCTCCAGGTGATGGTTTATTTCACCAITAGATACAGGCAA 1171
 QY 301 IleGlyProIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
 DB 1172 ATTGGTCCCTGATTGACAGGCAGATTTTCCGCTTTAGTGAAGATGGGATGCTAAATGCA 1231
 QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
 DB 1232 AGATTGACTATAGCTATGACACACAGCTTTCAGGTGACAGCATCAGGGTGTGATCAAT 1291

QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
 DB 1292 GAAACGCCACTGCCTATTGATCTGATCAGTTTGTGATGATTTCTGGCAAGTTTGGAGCAG 1351
 QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleThrThrAlaValMet 380
 DB 1352 TTTGAAAGTTTGGAGTTATATATATATATATTAACCAAGATCATTTCTACAGCTTAATG 1411
 QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
 DB 1412 ACCTATACGAGCACCTTTGATGCTCATGCGGATCAAGGATTCATATGATGATATTC 1471
 QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
 DB 1472 AGGTCGCTCATGCTACTGATTTACAATTCAGTATGATAACATGCGGTGCGGTAAACCAAGAGA 1531
 QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
 DB 1532 GAGATTAAATAGGCGCTTTTCCCAACACCAACCAATATGCTTATGAATATGATGTTGAT 1591
 QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
 DB 1592 GACACGCTCCAAACAGTTTACCTCAATGAAAGATAATGTGCGGTACCACTACCATCTG 1651
 QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
 DB 1652 AATGAAACCTCCATTTACTGAACCAAGTAAACAGTGCAGTGCAGTGCCTTCGCTAT 1711
 QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
 DB 1712 GACCTCGAGACAGAAATCACTCGACTGGGTGATGTTCAATATCGGTGGATGAAGATGGT 1771
 QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
 DB 1772 TTCTACGTCMAAGGCGCACGGAATCTTTGAATATAGCTCCAAAGGGGCTTCTAACTCGA 1831
 QY 521 AlaTyrAsnArgAlaGlySerThrSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
 DB 1832 GTTTACAGTAAAGGCGAGTGGCTGACAGTATCTACCGTTTATGACGCGCTGGGAAGCGGT 1891
 QY 541 Val 541
 DB 1892 GTT 1894

Search completed: August 14, 2004, 12:02:17
 Job time : 9668.56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 3555.05 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_2400_2600

Perfect score: 1077

Sequence: 1 IGYHGLYDPLTKLVHMR.....TDIISVANEDGRRVAAIILNH 201

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -qfmt=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US10029020 @CGN 1 1 19065 @runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_or.*
21: em_ov.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rnd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|-------|---------------------|
| 1 | 1077 | 100.0 | 3320 | 9 | AK127101 Homo sapi |
| 2 | 1077 | 100.0 | 3486 | 9 | BS537983 Homo sapi |
| 3 | 1077 | 100.0 | 8354 | 6 | AX556500 Sequence |
| 4 | 1077 | 100.0 | 8624 | 9 | AB037723 Homo sapi |
| 5 | 1077 | 100.0 | 8645 | 6 | AX600210 Sequence |
| 6 | 1077 | 100.0 | 8993 | 9 | BSM806812 Sequence |
| 7 | 1070 | 99.4 | 2685 | 9 | AK058531 Homo sapi |
| 8 | 1063 | 98.7 | 8438 | 6 | AK058531 Sequence |
| 9 | 1031 | 95.7 | 5583 | 10 | AK122490 Mus muscu |
| 10 | 1031 | 95.7 | 8585 | 10 | AB025413 Mus muscu |
| 11 | 1031 | 95.7 | 9722 | 10 | AF059485 Mus muscu |
| 12 | 807 | 74.9 | 9264 | 5 | AB026980 Danio rer |
| 13 | 667.5 | 62.0 | 8118 | 5 | GA038613 Gallus ga |
| 14 | 636.5 | 59.1 | 8373 | 10 | AB025410 Mus muscu |
| 15 | 630.5 | 58.1 | 8297 | 9 | AF100772 Homo sapi |
| 16 | 627 | 58.2 | 834 | 5 | GA039019 Gallus ga |
| 17 | 566 | 52.6 | 2157 | 6 | AX876360 Sequence |
| 18 | 566 | 52.6 | 2157 | 6 | BD156088 Primer fo |
| 19 | 566 | 52.6 | 2157 | 6 | AX877449 Sequence |
| 20 | 566 | 52.6 | 3270 | 6 | AK027473 Homo sapi |
| 21 | 566 | 52.6 | 3270 | 6 | BD156663 Primer fo |
| 22 | 566 | 52.6 | 3270 | 9 | AK001748 Homo sapi |
| 23 | 566 | 52.6 | 3614 | 6 | AX876525 Sequence |
| 24 | 566 | 52.6 | 3614 | 6 | BD156175 Primer fo |
| 25 | 566 | 52.6 | 3614 | 9 | AK001336 Homo sapi |
| 26 | 566 | 52.6 | 5309 | 9 | AB040888 Homo sapi |
| 27 | 566 | 52.6 | 8473 | 6 | AX662357 Sequence |
| 28 | 566 | 52.6 | 8487 | 6 | AX662359 Sequence |
| 29 | 566 | 52.6 | 8645 | 6 | AX662355 Sequence |
| 30 | 566 | 52.6 | 8675 | 6 | AX662353 Sequence |
| 31 | 563 | 52.3 | 5804 | 10 | AK122513 Mus muscu |
| 32 | 563 | 52.3 | 7816 | 10 | AF195418 Mus muscu |
| 33 | 563 | 52.3 | 8964 | 10 | AB025412 Mus muscu |
| 34 | 536.5 | 49.8 | 3184 | 10 | AF195419 Mus muscu |
| 35 | 536.5 | 49.8 | 8797 | 6 | AX250063 Sequence |
| 36 | 536.5 | 49.8 | 8797 | 6 | AX250066 Sequence |
| 37 | 536.5 | 49.8 | 8797 | 10 | AB025411 Mus muscu |
| 38 | 531.5 | 49.4 | 7713 | 9 | BSM808325 Sequence |
| 39 | 531.5 | 49.4 | 8689 | 6 | AX250067 Sequence |
| 40 | 531.5 | 49.4 | 8689 | 10 | AF086607 Rattus no |
| 41 | 529.5 | 49.2 | 2926 | 9 | BSM802230 Homo sapi |
| 42 | 529.5 | 49.2 | 6560 | 6 | AX250065 Sequence |
| 43 | 529.5 | 49.2 | 7781 | 9 | AB032953 Homo sapi |
| 44 | 529.5 | 49.2 | 9729 | 6 | AX250013 Sequence |
| 45 | 529.5 | 49.2 | 9826 | 6 | AX250008 Sequence |

ALIGNMENTS

RESULT 1

```

AK127101      3320 bp      mRNA      linear      PRI 09-SEP-2003
LOCUS      Homo sapiens cDNA FLJ45158 f1s, clone BRAWH3043034, highly similar
DEFINITION      to Mus musculus neuregulin 1 (Nrg1).
ACCESSION      AK127101
VERSION      AK127101.1 GI:34533862
KEYWORDS      oligo capping; f1s (full insert sequence).
SOURCE      Homo sapiens
ORGANISM      Homo sapiens (human)
REFERENCE      1
AUTHORS      Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J.,
Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K.,
Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H.,
Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K.,
Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
TITLE      UNPUBLISHED
JOURNAL
REFERENCE      2 (bases 1 to 3320)
AUTHORS      Isogai, T. and Yamamoto, J.
DIRECT SUBMISSION
SUBMITTED (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT      NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
Location/Qualifiers
1..3320
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRAWH3043034"
/tissue_type="brain"
/clone_lib="BRAWH3"
/note="cloning vector: pME18SFL3"
ORIGIN
Alignment Scores:
Pred. No.: 1.5e-111 Length: 3320
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_2400_2600 (1-201) x AK127101 (1-3320)
QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrIleValHisMetGlyArg 20
DB 1956 ATAGGCTACCATGGTGGCTTCTATGATCCACTCACCAGGCTGTCCACATGGCCGCGCA 2015
QY 21 AspTyrAspValLeuAlaGlyArgTTrpThrSerProAspHisGlyLeuTyrIleValHisLeu 40
DB 2016 GATTATGATGCTGTGGCGGACGCTGACTAGCCAGACACGAGCTGTGGAGCACCTT 2075
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheIleValAsnAsnProIleSer 60
DB 2076 AGTAGCAGCAACGTCATGCCCTTTTAAATCTCTATATGTTCAAAAACAACACCCCATCAGC 2135
QY 61 AsnSerGlnAspIleIleValHisPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
DB 2136 AACTCCAGGACATCAAGTGCCTTCATGACAGATGTTAAGAGCTGGCTGCTCCTTGGCA 2195
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProIleValProAspMetAspAlaMetGluPro 100

```

```

2196 TTCCAGCTACACAACGCTGATCCCTGGTTATCCCAACACAGACATGGATGCATGGAACCC 2255
QY 101 SerTyrGluLeuIleHisThrGlnMetIleThrGlnGlnTyrAspAsnSerIleSerIle 120
2256 TCCTACGAGCTCATCCACACACAGATGAAACGAGGAGTGGGCAACAGCAAGTCTATC 2315
QY 121 LeuGlyValGlnCysGluValGlnIleValGlnLeuIleValAlaPheValThrLeuGluArgPhe 140
2316 CTCGGGTACAGTGTGAAGTACAGACAGCTCAAGGCTTTGTACCTTAGAACGGTTT 2375
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProIleThrIleIleValPhe 160
2376 GACCAAGCTCTATGGCTCCACCAATCACCAGCTGCCAGCAGGCTCCAAACACCAAGAAGTTT 2435
QY 161 AlaSerSerGlySerValPheGlyValGlyValIleValPheAlaLeuLeuAspGlyArgVal 180
2436 GCATCCACGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCTTGAAGATGGCCGAGTG 2495
QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
2496 ACCACAGACATCATCAGTGTGGCCATCAGGATGGCGAAGGTTGCTGCCATCTTGAAC 2555
QY 201 His 201
2556 CAT 2558
RESULT 2
HSM806114
LOCUS      HSM806114
DEFINITION      Homo sapiens mRNA; cDNA DKFZp686D0412 (from clone DKFZp686D0412).
ACCESSION      BX537983
VERSION      BX537983.1 GI:31874053
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 3486)
AUTHORS      Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Newes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
TITLE      Direct Submission
JOURNAL      Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT      Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686D0412) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1..3486
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="11q14.1"
/clone="DKFZp686D0412"
/tissue_type="human cervix"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
1..3486
/gene="DKFZp686D0412"
<1..3217
/gene="DKFZp686D0412"
/note="hypothetical protein, N-terminus truncated"
/codon_start=2
/product="hypothetical protein"

```

```
/protein_id="CAD97943.1"
/db_xref="GI:31874054"
translation="SFGRLTNVFTTQGVSRSDTSSVHVQVETSKDDVTITNL
SASGAFYLLQDQVRNSYIIGADSLRLLANGMEVALQTEPHLLIAGTVNPTVGRNV
TLPLDINGLNLVEMQRKEARQGVTVFGRRLRVHNRLLSLGFDRTVTEKTYDDHRK
FLRLIYDQAGRPSLWSPSRNGVTVFGRRLRVHNRLLSLGFDRTVTEKTYDDHRK
IFADCKTWSYTYLEKSWLLHLSORQYIPEFDKNDRLSSVTMPNVAROTLETIRSVG
YRNTYQPPGASVITQDTEGHLHLLTFYLGATGRVVIYKXKLSKLAFTVDTTKVSF
TYTAGMLKTLINLQNEGFTCTIRYQIGPLIDRQIFRTEEGMVARFDYNDNSFR
VTSMAVNETPLPDLIRYDDVSKTEQFGKGVYIYDINOIITAVMTHTKHPDAY
GRMKEVQYIFRSLMTQVQVDMNRVVKKELKVPYANTRYSEYDADQQLQTVS
INDPLRYSYDGLNGLHLLSPGNSARLTPLRDIIRDRITRLGVOYQWMDGDFLQR
GGDIFEVNSAGLLIKAVNRAGSWSVRYDGLGRVSSKSSHLOPFYADLTNPYK
VTHLYNHSSRITSYLYDLOGLHFMELSSGDEFFIACDNIGTFLAVSGTGLMKQI
LYTAGELVMDTNEFNQIIGHGGLYDPLTKLVHMRDRYDVLGAGRWTSPPDLHWH
LSSNVMEFNLMFNPNISQDIKCFMTDVNSWLLTFGLFNLVNPYKDPMDMA
MEPSYELHTQMKTEWNSKSIILGQVCEQVQKLFVTLERFDLYGSTITSTCCQAP
KTKFPASSGVSFGDGLAIFGLSGRRRLTEGVNVTVOIINTVLNGTRRYTDIOLQY
GVDTHTFYKPGPSGDLAIFGLSGRRRLTEGVNVTVOIINTVLNGTRRYTDIOLQY
GALCLNTRYGTTLDEEAKRVLELQRAVRQAWAREQQLREGEGRLRWTEGKQQV
LSTGEVQGYDGFVVISVEQYFELSDSANNIHFMRQSENGRR"
3468
polya_site
/gene="DKFZp696D0412"

ORIGIN
Alignment Scores:
Pred. No.: 1.59e-111 Length: 3486
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x HSM806114 (1-3486)
QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
Db 2105 ATAGGCTACCATGCTGGCCCTATGATCCACTCACCAGCTGTGCCACATGGCGCGGA 2164
QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 2165 GATTATGATGTGCTGGCGGACGCTGGACTAGCCACAGCAGCAGCTGTGGAAGCACCTT 2224
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
Db 2225 AGTAGCAGCAGCAGCTGATGCTTTTAACTCTATATGTTTCAAAAACAACACCCCATCAGC 2284
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 2285 AACTCCAGGACATCAAGTGTCTCATCAGAGATGTTAAACAGCTGGCTGCTCACCTTTGGA 2344
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 2345 TTCAGCTACACACGATGATCCCTGCTTATCCAAAACAGACATGGATGCGTGAACCC 2404
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnLutrpAspAsnSerLysSerIle 120
Db 2405 TCCTAGAGCTTATCCACACAGATGAAACCGAGGATGGGACACACAGCAAGTCTATC 2464
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 2465 CTCGGGGTACAGTGTGAAGTACAGAGCAGCTCAAGGCCCTTTGTCACCTTAGAACCGTTT 2524
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 2525 GACACGCTCTATGGCTCCACAATCACCAGCTGCGAGCAGGCTCCAAAGACCAAGAGTTT 2584
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 2585 GCATCCAGCGGCTCAGTCTTTTGGCAGGGGGTCAAGTTTGCCTTGAAGGATGCCCGAGTG 2644
QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn 200
Db 7711 GACCACTCTATGGCTCCCAATCACAGCTGCCAGGCTCCAAAGACCAAGAGTTT 7711
```

Db 2645 ACCACAGACATCATCAGTGTGTGGCCAAATGAGGATGGCGGAGGGTTGTCTGCCATCTTGAAC 2704

QY 201 His 201
|||

Db 2705 CAT 2707
|||

RESULT 3
AX556500 8354 bp DNA linear PAT 27-NOV-2002
LOCUS AX556500
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
LOCUS Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Gangolli, E.A., Patturajan, M., Vernet, C.A., Malyankar, U.M.,
Keruda, R., Stone, D.J., Anderson, D., Shimkets, R.A., Burgess, C.E.,
Zerhusen, B.D., Liu, X., Spytek, K.A., Casman, S.J., Boldog, F.L.,
Smithson, G., Li, L. and Ji, W.
Polypetides and nucleic acids encoding same
Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
Location/Qualifiers
1..8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

TITLE
JOURNAL
FEATURES
source
1..8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4.52e-111 Length: 8354
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AX556500 (1-8354)

QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
Db 7232 ATAGGCTACCATGCTGGCCCTATGATCCACTCACCAGCTGTGCCACATGGCGCGGA 7291
QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 7292 GATTATGATGTGCTGGCGGACGCTGGACTAGCCACAGCAGCAGCTGTGGAAGCACCTT 7351
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
Db 7352 AGTAGCAGCAGCAGCTGATGCTTTTAACTCTATATGTTTCAAAAACAACACCCCATCAGC 7411
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7412 AACTCCAGGACATCAAGTGTCTCATCAGAGATGTTAAACAGCTGGCTGCTCACCTTTGGA 7471
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7472 TTCAGCTACACACGATGATCCCTGCTTATCCAAAACAGACATGGATGCGTGAACCC 7531
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnLutrpAspAsnSerLysSerIle 120
Db 7532 TCCTAGAGCTCTATCCACACAGATGAAACCGAGGATGGGACACACAGCAAGTCTATC 7591
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7592 CTCGGGGTACAGTGTGAAGTACAGAGCAGCTCAAGGCCCTTTGTCACCTTAGAACCGTTT 7651
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 7652 GACCACTCTATGGCTCCCAATCACAGCTGCCAGGCTCCAAAGACCAAGAGTTT 7711

Qy 161 AlasSerSerGlySerValPheGlyGlyVallyysPheAlaLeuLeuAspGlyArgVal 180
Db 7712 GCATCCAGCGCTCAGCTTTTGGCAAGGGGTCAAGTTTGCTTGAAGGATGGCGGAGTG 7771
Qy 181 ThrThrAspIleLeSerValAlaAsnGluAspGlyArgValAlaAlaLeuAsn 200
Db 7772 ACCACAGACATCATGAGTGGCCATGAGGATGGCGAAGGTTGCTGCCATCTTGAC 7831
Qy 201 His 201
Db 7832 CAT 7834
RESULT 4
AB037723 8624 bp mRNA linear PRI 10-MAY-2002
LOCUS Homo sapiens mRNA for KIAA1302 protein, partial cds.
DEFINITION
AB037723
ACCESSION
AB037723.2 GI:20521827
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1
Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
20181126
10718138
2 (bases 1 to 8624)
Ohara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnaifc@kazusa.or.jp.
URL: <http://www.kazusa.or.jp/huge/>, Tel:+81-438-52-3913,
Fax:+81-438-52-3914
On May 9, 2002 this sequence version replaced gi:7242958.
COMMENT
Location/Qualifiers
1. 8624
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="bf00104"
/tissue type="brain"
/clone_lib="pBCSKPGW"
/note="This sequence was replaced that of fg06864 cDNA as
a representative cDNA sequence for KIAA1302."
1. 8624
/gene="KIAA1302"
<590.3850
/gene="KIAA1302"
/note="Start codon is not identified."
/codon_start=1
/product="KIAA1302 protein"
/protein_id="BAA92540.2"
/db_xref="GI:20521828"
/translation="STERLDNKPETTRYDSFGRLTNVTPTQGVSSFRSDTSSVHVQ
VETSSKDDVTITNLGASGAFYTLQDOVRNSYVIGADGSLRLLLANGMEVALQTEPH
LLAGTVPVTKENVTLPIDNGLNLMVQRKEQARGQVTFGRRLRVHNRNLLSLDF
DVRTKEKYDDRRKTELRLYDQACRPSLWSPSSRLNGVNTVYSGGYTAGIQRGIM
SERMEYDQAGRTSRIFADGKTSYITLKSQVLSVQDFEDGHLHLTYLFGTGRVRYTKGL
NVARQLETIRSYGYNTYQPPGNASVIQDFEDGHLHLTYLFGTGRVRYTKGL
SKLAETLYDTKVSFYDTAGMKLTINLQNEGFTCTIRYQIGLIDQIFRFTGEG
MNVARPDYNDNFRVTSQAVINETPLIDLYRVDYDVSQKTFQKFGVYVYDINQI
ITVAVMTHKHPDAYGRMKVEQVEIFRSLMVMVQVNDNMVVKELKVPYANITR
YSVEYDQGLQTVSYNDKPLWEYSYDNLGNLHLLSPGNSARLTPLRYDIRDRITRLG
DYQYKMDGFLQRQGDIFEYNSAGLLIKAYNAGSWSVRYRDGLGRVSSKSHS
HLQFYADLTPTKTYLHNSSEITSLYDLQGLHFAWELSSGDEFYIACDNTGT
PLAVFSGTGLMIKQILYATYGEIYMDTNFNFIIGYHGLGYDPLTKLVHMGRRDYDV

LAGRWTSDEHLMKHLSSNMPFNLMKNNNPISNSQDIKCFMTDVNSWLLTFGFG
LHVIPGYPPDMDAMEPSYELIHTOMKTQEWDSKSLGVCQEVQKAKFVTLERF
DOYLSGTTSCQAPKTKGSKVFKGKVFALQKGRVTTDIISVANEDGRVAAI
LNHAHYLENLHFTTIDGVTYFVKPSPSEGLAILGSGRRITLNGVNVTSQINTV
LNGTRRYTDLQYQALCINTRYGTTLDEKARVLELQRAVROQAWARQORLREG
BEGURAWTEGEKQVQLSTGRVQVGGFFVISVEQYPELSDSANNIHFMRQSEMER"
ORIGIN
Alignment Scores:
Pred. No.: 4,69e-111 Length: 8624
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_2400_2600 (1-201) x AB037723 (1-8624)
Qy 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 2738 ATAGGCTACCATGTTGGCTCTTATGATCCACTCACAAGCTTGTCCACATGGCGCGGA 2797
Qy 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 2798 GATTATGATGCTGCTGCCGAGCGCTGGACTAGCCAGACCCAGAGCTGTGGAGACCTT 2857
Qy 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProLeSer 60
Db 2858 AGTACAGCAAGCTCATGCCCTTTTATCTCTATATGTTCAAAAACAACACCCCATCAGC 2917
Qy 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrPleuLeuThrPheGly 80
Db 2918 AACTCCAGGACATCAAGTGTCTCATGACAGATGTTAAACAGCTGGCTGCCTTTGGA 2977
Qy 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 2978 TTCCAGCTACACAACAGTATCCCTGTTATCCCAACACAGACATGATGCCATGGNACC 3037
Qy 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
Db 3038 TCTTACGAGCTTATCCACACAGATGAAACGAGGAGTGGGACACACAGCAAGTCTATC 3097
Qy 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 3098 CTGCGGGTACAGTGTGAGTACAGAGCAGCTCAAGGGCTTTTGTCACTTAGAACCGTTT 3157
Qy 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 3158 GACCACTCTATGGCTTCCAAATCACCAGCTGCCAGAGCTCCAAAGACCAAGAGTTT 3217
Qy 161 AlasSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 3218 GCATCCAGCGGCTCAGCTTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCGGAGTG 3277
Qy 181 ThrThrAspIleLeSerValAlaAsnGluAspGlyArgValAlaAlaLeuAsn 200
Db 3278 ACCACAGACATCATGAGTGGCCATGAGGATGGCGAAGGTTGCTGCCATCTTGAC 3337
Qy 201 His 201
Db 3338 CAT 3340
RESULT 5
AX600210 8645 bp DNA linear PAT 14-FEB-2003
LOCUS Sequence 22 from Patent WO02072830.
DEFINITION
AX600210
ACCESSION
AX600210.1 GI:28400252
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Yue, H., Yao, M.G., Ison, C.H., Lu, Y., Warren, B.A., Elliott, V.S.,
Baughn, M.R., Ding, L., Xu, Y., Gietzen, K.J., Tang, T.Y., Lal, P.G.,
Duggan, B.M., Burford, N., Lu, D.A., Richardson, T.W., Tran, U.K.,
Khare, R. and Wallia, N.K.
TITLE Proteins associated with cell growth, differentiation, and death
JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
Incyte Genomics, Inc. (US)
FEATURES
source 1..8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7488573CB1"
ORIGIN
Alignment Scores:
Pred. No.: 4,71e-111 Length: 8645
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_2400_2600 (1-201) x AX600210 (1-8645)
QY 1 IleglyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7281 ATAGGCTACCATGCTGGCTCTATGATCCACTACCAAGCTGTGCCAGCGCGCA 7340
QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 7341 GATTATGATGTGCTGGCGGAGCTGACTAGCCACACACGAGCTGTGGAAGCACCTT 7400
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7401 AGTAGCAGCAACGTCATGCGCTTTTAACTCTATATGTTCAAAACACCAACCCATCAGC 7460
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuThrPheGly 80
Db 7461 AACTCCAGGACATCAAGTCTCATGACAGATGTTAAACAGCTGGCTGCCTTTGGA 7520
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7521 TTCAGCTACACACAGTATCCCTGCTATCCCAACACAGACATGATGCATGGAACCC 7580
QY 101 SerTyrGluLeuLeuHisThrGlnMetLysThrGlnLysTrpAspAsnSerLysSerIle 120
Db 7581 TCTACAGCTCATCCACACACAGATGAAACGAGAGTGGGACACAGCAAGTCTATC 7640
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7641 CTCGGGTACAGTGTGAGGTACAGAGCAGCTCAAGCCCTTGTCCACCTTAGAACGGTTT 7700
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 7701 GACCAGCTCATGGCTCCCAATCACCAGCTGCCAGCGCTCCAAAGACCAAGAGTTT 7760
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 7761 GCATCAGCGGCTCAGTCTTGTGCAAGGGGGTCAAGTTTGCCTTGAAGGATGCCGAGTG 7820
QY 181 ThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
Db 7821 ACCACAGATCATCATGCTGTGCCAATGAGGATGGCGGAGGGTGTGCTCCCTTCTGAAC 7880
QY 201 His 201
Db 7881 CAT 7883
RESULT 6
LOCUS HSM806812 8993 bp mRNA linear PRI 28-AUG-2003

DEFINITION Homo sapiens mRNA; cDNA DKFZp686K11107 (from clone DKFZp686K11107).
ACCESSION BX640737
VERSION BX640737.1 GI:34364828
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 8993)
AUTHORS Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
CONSRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686K11107) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/
FEATURES
source 1..8993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="11q14.1"
/clone="DKFZp686K11107"
/tissue type="human fetal kidney"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="fetal"
gene 1..8993
/genes="DKFZp686K11107"
CDS <1..4200
/gene="DKFZp686K11107"
/note="hypothetical protein, N-terminus elongated,
differentially spliced"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA45850.1"
/db_xref="GI:34364829"
/translation="RIDQNGIISTLLGSDLTSLRPLSCDSVNDISQVHLEWPTDLAI
LAVSHNGVLIYIAETDEKKNIRIQVTSGEISLVAGAPGCCCKNDANDCFSGDDGY
AKDLNTPSSSLAVCADGELYVADLGNIRIRIRKNKPLNTQNMVELSSPIDQELYL
FDTTGKHLVTSQSLPTGDLNFTYTGDDGDTLTIDNNGMNVNRDSTCGMLVLPD
GOVYVMTGNTSALKSVTTQGHLEAMVTYHNGSGLLATKSNENGWTFYEDSFGRT
NVTFTGOVSSEFSDTSVHVQVETSSKDDVTITNLASGAFYTLTLDQVRNSYII
GADGSRLLLANGMEVALQTEPHLLAGTPTVGRKNTLPIDNGLNLEWQRKQQA
RGQTVFGRRLRHNRLSLDFRTEKTYDDHKEFTLRILYDQAGRPGLSWSPSS
RLNGVNTYSPGGYTAGIQGIRMSRMEYDQAGRTSRIFADKWTWSYVLEKSMVL
LHQQRYTFEFGNDRLSSVTMPNVAROTLETIRSVGYVYRNYQPEGNASVTDQTE
DGHQRYFLGTGRVIVYKGLSKLAETLYDTTKVSFYDSTAGMLKTNLQNEGT
CTIRYQIGPLIDRQIFPTEGWNARFDINDNSFRVTSQAWKTEYFISLWMYTV
DVDSKQFGKFGVYIYDINQIITAVMHTKHFDAVGRMEVQVYFISLWMYTV
QYDNKRVYKGLKVGPNATRYSDYDADQQLQTVSINDKPLFYSIDNGLNHL
SPGNSARLTPLRDRITRLGQVYKMDGDFLRQGGDI FEVNSAGLLIKAYNRA
GSWSVRYRDTGLGRVRSKSSHSHLQFPYADLTNPYKVTLYLHNSSEISLTYDLQ
GHLFAMELSSGDEFYIACDNIQTPLAVFSGTGLMIKQLIYATGYIDTDNFIQII
GHLGGLDPLTKLVHMGRRDIDLAVRWTPSPHLEWKLSSNVNMFNLMFNKNPI
SNSDIKCFMTDINWLLTGFQLEHVTIPGYKPDMDAMEPSYELIHTQMTQWONS
KSILGVQCEVQKQLAFVTLERFDQLYGSTITSQQAPKTKFKFASGSGVKGKVPAL
KDGRTVTDIIISVANEDGRVAAILNHAHLENLHFTIDGVDTHYFVKFPGSGDAIL
GLSGGRTLENGVNVTSQINTVLNGRTRRYTDIOLOYGALCLNRYGTILDEKARV
LELAFORAVROAWAREQORLEBEGELRAWTEGEKQQVLSIGRVQGYDGFVFSVEQY
PELSDSANNIHFMRQSEMR" 8953..8958
polyA_signal

```

polyA_site      /gene="DKFZp686K11107"
8973
/gene="DKFZp686K11107"

ORIGIN
Alignment Scores:      4.94e-111      Length:      8993
Pred. No.:      1077.00      Matches:      201
Score:      100.00%      Conservative:      0
Percent Similarity:      100.00%      Mismatches:      0
Best Local Similarity:      100.00%      Indels:      0
Query Match:      9      Gaps:      0
DB:

US-10-029-020-14_COPY_2400_2600 (1-201) x HSM806812 (1-8993)

Qy      1      IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
Db      3088      ATAGGCTACCATGGTGGCTCTATGATCCACTCACCAGCTGTGCCATGGCGCGGA 3147

Qy      21      AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db      3148      GATTATGATGTCTGGCGCGACGCTGGACTAGCCAGACCGAGCTGTGGAAGCACCTT 3207

Qy      41      SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db      3208      AGTAGCAGCAACGTCATGCTTTTAACTCTATATGTTCAAAACAACAACCCCATCAGC 3267

Qy      61      AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrLeuLeuThrPheGly 80
Db      3268      AACTCCAGGACATCAAGTGTTCATGACAGATGTTAAAGCTGGCTGTCTACCTTTGGA 3327

Qy      81      PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db      3328      TTCCAGCTACACACGTCATGCTGCTTATCCCAACACAGACATGATGCCATGAACCC 3387

Qy      101      SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
Db      3388      TCTTACGAGCTTATCCACACAGATGAAACGCGAGGAGTGGCAACAGCAAGTCTATC 3447

Qy      121      LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db      3448      CTCGGGTACAGTGTGAAGTACAGACAGAGCTCAAGGCTTTTGTACCTTAGAAGCGTT 3507

Qy      141      AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db      3508      GACCAGCTCTATGGCTCCCAATCACATCACAGTCCAGCTGCCAGAGCTCCAAAGACCAAGAGTTT 3567

Qy      161      AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db      3568      GCATCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGCCGAGTG 3627

Qy      181      ThrThrAspIleLeuSerValAlaAsnGluAspGlyArgArgValAlaIleLeuAsn 200
Db      3628      ACCACAGACATCATCAGTGTGGCAATGAGGATGGCGGAGGGTGTCTGCCATCTTGAAC 3687

Qy      201      His 201
Db      3688      CAT 3690

RESULT 7
AK056531
LOCUS      Homo sapiens cDNA FLJ31969 fis clone NT2RP7008013, highly similar
to Mus musculus mRNA for Ten-m4.
DEFINITION
AK056531
ACCESSION      AK056531.1 GI:16551956
VERSION
KEYWORDS      oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Theria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS      Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
Unpublished
NEDO human cDNA sequencing project
2 (bases 1 to 2685)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
JOURNAL
TITLE
JOURNAL
AUTHORS
REFERENCE
TITLE
JOURNAL
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.
FEATURES
source
Location/Qualifiers
1..2685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP7008013"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP7"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 5-weeks retinoic acid (RA)
induction."
242..2434
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71206.1"
/db_xref="GI:16551957"
/translation="MLKTNLQNEGFTCTIRYQIGPLIDROI FRFTBGMVNARFDY
TKHFDAYGRMEVQVEIFRSLMWTVDNNRVVKELKGVKVPVANTIRYSYEDAD
GQLQVSNIDKFLMYSDLNLGNLHLLSPGNSARLTPLURIDRITRUGDVQYKMD
DGLFQRGDGIPEYNSAGLLIKAYNRAGSWSVRYDGLGRVSKSSSHHLPFFYA
DLTNPKTHLYNHSSSTTSLYDLQHLFAMELSSGDEFIACDNICTPLAVFSGT
GLMIQLIYATGEIYMTNPNFQIIIGYVHGQIDPLTKLVHMRDRDYDLAGRWTSP
DHELWKRHSSNNVFNLFYKNNPINSQDIKCEMTDVSMLITFGQLHNVIPGY
PKPDMAMEPSYELIHTQMTQEDWNSKSLGVQCEVQKLFATVLESFDOLYGSTI
TSCLOAPKTKFPASSGVFGKGVKFDALDGRVTDTISVANEDGRVVAALLNHAHYLE
NLHFTIDGVDTHTYFKPGPSEGDLAILGLSGGRRLTLENGVNTVTSQINTVLSGRTHRY
TDIQLQYGCALCLNTRYGTTLDEKARVLELRQAVRQAWAREQQRLEGEGLRAWT
EGEKQQLVSTGRVQGVDFVSVQYPELSDSANNIHFMRQSEMGRR"
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71206.1"
/db_xref="GI:16551957"
/translation="MLKTNLQNEGFTCTIRYQIGPLIDROI FRFTBGMVNARFDY
TKHFDAYGRMEVQVEIFRSLMWTVDNNRVVKELKGVKVPVANTIRYSYEDAD
GQLQVSNIDKFLMYSDLNLGNLHLLSPGNSARLTPLURIDRITRUGDVQYKMD
DGLFQRGDGIPEYNSAGLLIKAYNRAGSWSVRYDGLGRVSKSSSHHLPFFYA
DLTNPKTHLYNHSSSTTSLYDLQHLFAMELSSGDEFIACDNICTPLAVFSGT
GLMIQLIYATGEIYMTNPNFQIIIGYVHGQIDPLTKLVHMRDRDYDLAGRWTSP
DHELWKRHSSNNVFNLFYKNNPINSQDIKCEMTDVSMLITFGQLHNVIPGY
PKPDMAMEPSYELIHTQMTQEDWNSKSLGVQCEVQKLFATVLESFDOLYGSTI
TSCLOAPKTKFPASSGVFGKGVKFDALDGRVTDTISVANEDGRVVAALLNHAHYLE
NLHFTIDGVDTHTYFKPGPSEGDLAILGLSGGRRLTLENGVNTVTSQINTVLSGRTHRY
TDIQLQYGCALCLNTRYGTTLDEKARVLELRQAVRQAWAREQQRLEGEGLRAWT
EGEKQQLVSTGRVQGVDFVSVQYPELSDSANNIHFMRQSEMGRR"
ORIGIN
Alignment Scores:      7.2e-111      Length:      2685
Pred. No.:      1070.00      Matches:      200
Score:      99.50%      Conservative:      0
Percent Similarity:      99.50%      Mismatches:      1
Best Local Similarity:      99.50%      Indels:      0
Query Match:      99.35%      Gaps:      0
DB:

US-10-029-020-14_COPY_2400_2600 (1-201) x AK056531 (1-2685)

Qy      1      IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
Db      1322      ATAGGCTACCATGGTGGCTCTATGATCCACTCACCAGCTGTGCCATGGCGCGGA 1381

Qy      21      AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db      1382      GATTATGATGTCTGGCGCGACGCTGGACTAGCCAGACCCAGACCCAGCTGTGGAAGCACCTT 1441

```

| | | | |
|----|------|---|------|
| QY | 41 | SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer | 60 |
| Db | 1442 | AGTAGCAGCACGTCATGCCCTTTAAATCTCTATATGTTCAAAAACAACACCCCATCAGC | 1501 |
| QY | 61 | AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrLeuLeuThrPheGly | 80 |
| Db | 1502 | AACTCCCAGGACATCAAGTGCTTCATGACACATGTTAAACAGCTGGCTGCTCACCTTTGGA | 1561 |
| QY | 81 | PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro | 100 |
| Db | 1562 | TTCAGCTTACACAACGTCATCCCTGGTTATCCCAAACGACATGGATGCCATGAAACC | 1621 |
| QY | 101 | SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle | 120 |
| Db | 1622 | TCCTACGAGCTTATCCACACACATGAAACGCGAGGTGGGCAACAGCAAGTCTATC | 1681 |
| QY | 121 | LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluAtrPhe | 140 |
| Db | 1682 | CTCGGGTCAGTGTGAAGTACAGACAGCTCAAGGCTTTGTCACTTAGAACGGTTT | 1741 |
| QY | 141 | AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe | 160 |
| Db | 1742 | GACCAGCTCTATGGCTCCCAATCACCAGTGCCTCGCAGGCTCCAAAGACCAAGAAGTTT | 1801 |
| QY | 161 | AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal | 180 |
| Db | 1802 | GCATCCAGCGGCTCAGTCTTTGGCAAGGGGGTCAAGTTTCCTTGAAGGATGGCCGAGTG | 1861 |
| QY | 181 | ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaIleLeuAsn | 200 |
| Db | 1862 | ACCACAGACATCATCAGTGTGGCCATCAGATGGCGAAGGGTGTGCTGCATCTTTGAAC | 1921 |
| QY | 201 | His | 201 |
| Db | 1922 | CAT | 1924 |

| | | | | | |
|------------|--|-------------|-----|--------|-----------------|
| RESULT 8 | AX675551 | 8438 bp | DNA | linear | PAT 27-MAR-2003 |
| LOCUS | AX675551 | | | | |
| DEFINITION | Sequence 1 from Patent WO02055704. | | | | |
| ACCESSION | AX675551 | | | | |
| VERSION | AX675551.1 | GI:29333352 | | | |
| KEYWORDS | | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| REFERENCE | 1 | | | | |
| AUTHORS | Padigaru, M., Li, L., Zerhusen, B.D., Casman, S.J., Shenoy, S., Spytek, K.A., Zhong, M., Ganggilli, E.A., Burgess, C.E., Patturajan, M., Vernet, C.A., Taylor, S., Tchernev, V.T., Miller, C.E., Guo, X., Balogh, F.L., Grosse, W.M., Alsbrook, J.P., Gerlach, V., Edingermark, S., Rothenberg, M.E., Ellerman, K., Macdougall, J., Malyankar, U., Millet, I., Peyman, J., Smithson, G., Gunther, E. and Stone, D.J. | | | | |
| | Proteins, polynucleotides encoding them and methods of using the same | | | | |
| TITLE | | | | | |
| JOURNAL | Patent: WO 02055704-A 1 18-JUL-2002; | | | | |

| DB: | 6 | Gaps: | 1 |
|---|--|---|-----------------------------|
| US-10-029-020-14_COPY_2400_2600 (1-201) x AX675551 (1-8438) | | | |
| QY | 1 | IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg | 20 |
| Db | 7259 | ATAGCTACCAATGTTGGCCCTCTATGATCCACTCACCAGCTTGTCCACATGGCGCGCGA | 7318 |
| QY | 21 | AspTyrAspValLeuAlaGlyArgTTrpSerProAspHisGluLeuTprLysHisLeu | 40 |
| Db | 7319 | GATTATGATGTGCTGGCGGACGCTGACTAGCCAGACCACGAGCTGTGGAGACCTT | 7378 |
| QY | 41 | SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer | 60 |
| Db | 7379 | AGTAGCAGCAACGTCATGCCTTTTAATCTCTATATGTTCAAAAACAACACCCCATCAGC | 7438 |
| QY | 61 | AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly | 80 |
| Db | 7439 | AACTCCCGAGACATCAAGTGCCTTCATGACAGATGTTAAACAGCTGGTGTCTACCTTTGGA | 7498 |
| QY | 81 | PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro | 100 |
| Db | 7499 | TTCCAGCTACACAACGTGATCCCTGGTTATCCCAAAACAGACATGATGCCATGGAAACCC | 7558 |
| QY | 101 | SerTyrGlnLeuIleHisThrGlnMetLysThrGlnGluTprAspAsnSerLys | 118 |
| Db | 7559 | TCCTACAGCTCATCACACACAGATGAAACCGCAGAGTGGGACACCAAGAGGTAAAT | 7618 |
| QY | 119 | -----SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLys | 132 |
| Db | 7619 | CTGCGACAAGGTCGCAGTCTATCTCTGGGGTACAGTGTCAAGTACAGAAGCAGCTCAAG | 7678 |
| QY | 133 | AlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGln | 152 |
| Db | 7679 | GCCTTTGTCCACTTAGAACGGTTTGACCACTCTATGGCTCCAAATCACCAGTGCACG | 7738 |
| QY | 153 | GlnAlaProLysThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyValLys | 172 |
| Db | 7739 | CAGGCTCCAAAGACCAAGAAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAG | 7798 |
| QY | 173 | PheAlaLeuLysAspGlyArgValThrThrAspIleIleSerValAlaAsnGluAspGly | 192 |
| Db | 7799 | TTTGCCCTTGAAGATGGCGAGTGACACAGACATCATCAGTGTGGCCAATGAGGATGG | 7858 |
| QY | 193 | ArgArgValAlaAlaIleLeuAsnHis | 201 |
| Db | 7859 | CGAAGGTTGCTGCCATCTTGAACCAT | 7885 |
| RESULT 9 | | | |
| LOCUS | AK122490 | 5583 bp | mRNA linear ROD 15-MAR-2003 |
| DEFINITION | Mus musculus mRNA for mKIAA1302 protein. | | |
| ACCESSION | AK122490 | | |
| VERSION | AK122490.1 GI:28972711 | | |
| KEYWORDS | FLI_CDNA. | | |
| SOURCE | Mus musculus (house mouse) | | |
| ORGANISM | Mus musculus | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| AUTHORS | 1 Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S., Nakajima,D., Nagase,T., Ohara,O. and Koga,H. | | |
| TITLE | Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries | | |
| JOURNAL | DNA Res. 10, 35-48 (2003) | | |
| REFERENCE | 2 (bases 1 to 5583) | | |
| AUTHORS | Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan | | |

(E-mail: mouse@kazusa.or.jp, Tel. 81-438-52-3919, Fax: 81-438-52-3918)
The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing: Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.

FEATURES

source
1. .5583
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mbg04788"
/tissue_type="brain"
/dev_stage="adult"
/note="vector:modified pBC SK+"
1. .5583
/gene="mKIAA1302"
1. .3597
/gene="mKIAA1302"
/note="CDS is predicted by in silico analysis. Start codon
is not identified."
/evidence="not experimental"
/product="mKIAA1302 protein"
/protein_id="BAC65772.1"
/db_xref="GI:28972712"
/translation="TQMYELSSPIDOELYLFDTSKHLXTQSLPTGDLVNFYVTGD
GDITHTDNNMNVNRDSTGMELMLVPPQVYVMTGNTSALRSVTTQGHLEAMV
TYHNSGLLATKSNENGWTFYEDSFGRLNVTPTQVSSFRSDTSSVHVQVETS
SKDDVTITNLSASGAFYLLQDQVRNSYIIGADSLRLLLLANGMEVALQTEPHLLAG
TVNPTGVRNVTLPIDGNLNVQRKEQVQVTFVRRRLVHNRNLLSLDFDRVT
ETEKYDHRKFTIRLLDQAGPSLSPSSRLNGVNVYSPGGHIAQIGQIMSEMR
EYDQGRITSIFADGKMSVYLEKSMVLHLHSORQVIFEDKNDRLSSVTPMNVLA
QTLRTSVYIRNIIQPEGNASVIOFTEDGHLHFTYLGTRVLYKYLKSLA
ETLDTTVTSFTYTDGMLKTNVQNEGFTCTIRYQIGLIDIDKQIFRTEBGMVNA
REFYNDMSFRVTSQAVINETPLIDLYRYDDVSGTEQFGKFGVYIYDINQIITTA
VMTQKEDVGRMEQVETPRSLMYWMTQVDMNRVVKELVGYANTRYSYE
YADQLOVTSINDKPLWYSYDINGNLLHLLSPGNSARLTPRLYDLRLRITRLGDOVY
KMDGDFRQRGDVFENSGALLIKAYNRASGWSVRYDGLGRVRSKSHSHLO
FFYADLTNPKTVHLYNHSSEITSLSYDLOGLHFAELSSGDEYIACDNIGTPLAV
FSGTGLMKTQYLYHNSSEITSLSYDLOGLHFAELSSGDEYIACDNIGTPLAV
WTSPPHEWLKRLSSNIVFPHLYMKNPNISNQIDKCFMTDVNSMLTFFQFQHN
IPGVPKPTDAMEPEYSLVHTQMKTEQWNSKSLGVOCEQOKLQAFVTLERFQPLY
GSLTNSCOQAPTEKPFASGSIFFGKGVFALKDGRVTTDIISVANEDGRRITAAIILNA
HYLENHETIDGVDHYFVKPGPSGDLAILGLSGRRTLENGVNVYVQINMLSGR
TRRYTDIQYALCINTRYTGVTEBEKRVLELQRAVRQAMAREOQLREGEGL
RAWTDGKQQLVNTGRVQYDGFVTSVEQYFELSDSANNHFMKQSENGR"

gene

CDS
1. .3597
/gene="mKIAA1302"
/note="CDS is predicted by in silico analysis. Start codon
is not identified."
/evidence="not experimental"
/product="mKIAA1302 protein"
/protein_id="BAC65772.1"
/db_xref="GI:28972712"
/translation="TQMYELSSPIDOELYLFDTSKHLXTQSLPTGDLVNFYVTGD
GDITHTDNNMNVNRDSTGMELMLVPPQVYVMTGNTSALRSVTTQGHLEAMV
TYHNSGLLATKSNENGWTFYEDSFGRLNVTPTQVSSFRSDTSSVHVQVETS
SKDDVTITNLSASGAFYLLQDQVRNSYIIGADSLRLLLLANGMEVALQTEPHLLAG
TVNPTGVRNVTLPIDGNLNVQRKEQVQVTFVRRRLVHNRNLLSLDFDRVT
ETEKYDHRKFTIRLLDQAGPSLSPSSRLNGVNVYSPGGHIAQIGQIMSEMR
EYDQGRITSIFADGKMSVYLEKSMVLHLHSORQVIFEDKNDRLSSVTPMNVLA
QTLRTSVYIRNIIQPEGNASVIOFTEDGHLHFTYLGTRVLYKYLKSLA
ETLDTTVTSFTYTDGMLKTNVQNEGFTCTIRYQIGLIDIDKQIFRTEBGMVNA
REFYNDMSFRVTSQAVINETPLIDLYRYDDVSGTEQFGKFGVYIYDINQIITTA
VMTQKEDVGRMEQVETPRSLMYWMTQVDMNRVVKELVGYANTRYSYE
YADQLOVTSINDKPLWYSYDINGNLLHLLSPGNSARLTPRLYDLRLRITRLGDOVY
KMDGDFRQRGDVFENSGALLIKAYNRASGWSVRYDGLGRVRSKSHSHLO
FFYADLTNPKTVHLYNHSSEITSLSYDLOGLHFAELSSGDEYIACDNIGTPLAV
FSGTGLMKTQYLYHNSSEITSLSYDLOGLHFAELSSGDEYIACDNIGTPLAV
WTSPPHEWLKRLSSNIVFPHLYMKNPNISNQIDKCFMTDVNSMLTFFQFQHN
IPGVPKPTDAMEPEYSLVHTQMKTEQWNSKSLGVOCEQOKLQAFVTLERFQPLY
GSLTNSCOQAPTEKPFASGSIFFGKGVFALKDGRVTTDIISVANEDGRRITAAIILNA
HYLENHETIDGVDHYFVKPGPSGDLAILGLSGRRTLENGVNVYVQINMLSGR
TRRYTDIQYALCINTRYTGVTEBEKRVLELQRAVRQAMAREOQLREGEGL
RAWTDGKQQLVNTGRVQYDGFVTSVEQYFELSDSANNHFMKQSENGR"

ORIGIN

Alignment Scores:
Pred. No.: 4,49e-106 Length: 5583
Score: 1031.00 Matches: 189
Percent Similarity: 99.00% Conservative: 10
Best Local Similarity: 94.03% Mismatches: 2
Query Match: 95.73% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AK122490 (1-5583)

QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThryLysLeuValHisMetGlyArgArg 20
Db 2485 ATCGGCTACACGGCGGCTCTATGATCCACTCACCAGGCTTGTCCACATGGCGCCACGG 2544
QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGlyLeuTyrLysHisLeu 40
Db 2545 GATTATGATGTGCTGGACGCTGGACAGCCCGACACCATGACTCTCGAAGCGCTG 2604
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
Db 2605 AGTAGCAACAGCATCGTGGCTTTTCATCTCTACATGTTTAAAGAACACACCCCATCAGC 2664
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuThryPheGly 80
Db 2665 AACTCTCAGGACATCAAGTCTTCATGACAGATGTCAACAGCTGGCTCCCTCACCTTTTGA 2724

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 2725 TTCAGCTGCACACGCTGATACCTGGCTATCCCAAGCCACACAGATGCTGGAACCC 2784
QY 101 SerTyrGlnLeuLeuHisThrGlnMetLysThrGlnLeuTyrAspAsnSerLysSerIle 120
Db 2785 TCCTACGAGCTCGTACACACACACATGAAACTCAGAAATGGGCAACAGCAAGCTCATC 2844
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 2845 CTCGGGCTACAGTGTGAAGTTCAGAACCACTCAAGCTTTCGTTACCTTAGAAGCTTT 2904
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThryLysLysPhe 160
Db 2905 GACCAGCTCTACGCTCGACCATCACCAGCTGCTCAACAGAGCCCTGAGCAAAAGATTT 2964
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 2965 GCCTCCAGTGTCTCCATCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGATGGTCGAGTG 3024
QY 181 ThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
Db 3025 ACCACTGACATCATCAGTGTGCGCAATGAGGATGGCGGAGGATCGCAGCCATCTTGAAC 3084
QY 201 His 201
Db 3085 AAT 3087

RESULT 10
AB025413
Mus musculus mRNA for Ten-m4, complete cds.
LOCUS
DEFINITION
AB025413
AB025413.1 GI:4760781
VERSION
KEYWORDS
Ten-m4.
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Ohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T.,
Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
J. Cell Biol. (1999) In press
2 (bases 1 to 8585)
Ohashi,T.
Direct Submission
Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
(E-mail: ohashi@cc.okayama-u.ac.jp, Tel.: +81-86-235-7128,
Fax: +81-86-222-7768)
FEATURES
Location/Qualifiers
1. .8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1. .8585
/gene="ten-m4"
188. .8503
/genes="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"
gene
CDS
188. .8503
/translation="MDVKRKPYSLTRRDAERRTSSADSEEGKPGKQKSSSET
LKAYDQDARLAYGRVKMDVMPQAEAFRCRTGNTFLRELGLMTPPHGLTYRTDGL
PHOGYSMGASSDADLEADTVLSPHPVLRWGRSTRSGRSSCLSRANSNLTLTDTEHE

NTETDHPSSQLQHPRLRTPPLPHAHNTNQHHAASINSINLRGNFTPRSPAPATDTH
 SLSEPPAGSAGEPETHAQONMLNINSPILETSLNKGKOPFLGTLQONLIEMDLISASRH
 DAGYSDGHLFLFKPGSTPCTPCTSPCTVSPSPRPPLPLPRSTPSPAPNLKSKPS
 KYCNWKAALISATILVILLAFVAMHLEGLNWHLOPMEGQWQMTIEDTASSW
 PVTDLVSLPSGGTGLETFDRKGAAGKSPSLFPEDSIDSEIDVGRRASQKIPF
 GTFWRSQVIFDHPHLEFNVSJGKAALVIGYRKGKLPSPSTQDFVELLDGRRLITQE
 ARSLGPGQKSGRPPSPSHETGFTQYLDGSIWHILAFYNDGKESVPLTVAIESVD
 NCPSYNGNDCISGTHCFGLGDCGRASCPVLSCNGQYMGKRCILCHSGMAE
 CDVPTNCIDVASCSSHGTCIMGTCTCNPGYKGSCEEVDCMDPTCCSRGVCVRGCHC
 SVWGTCNCTPRATCDQCSGHGTFLEPDTGLCNCDPSTWHDCEIEICADCCGHV
 CVGTCRCDGWMGACDORACHPRCAEHGTCDKGCSCSPGMEHCHTIAHLDVV
 KEGCPLNGNRCITLDLNGHVCOLGWRGTGCDTSMETGCGDKNDGDLVDCMD
 PDCCQLPCHVNPGLSPGPDLDIIETQAPVQSQNLNPFYDRIKFLVRDSTHISIPG
 ENPFDGHCACVIRGQVSTDPGLVFNISFINNPLFGVTISRODPSFDLVTNGGISI
 ILRFRAPFTIQEHTLWLPDWRFFVMEITVMEHENEIPSCDLSNFPANPVPSPPL
 TSFASCAEKGPIVEICAOLOBEIIVAGCKMRLSLSSRTPGKYSRLRSLTHPTPF
 NLMKVLHMAVEGRFLFRWFAAAPDLSYFIWDKTDVYNQKVFGESEAFVSVEYES
 CPDLLLMEKRTAVLGQIEIDASKGUGSLDKHNLNIOGILHKGNGENQFVSQPPV
 IGSIMGNRRRSISCPSCNGLADGNKLPAVALTCGSDGSLYGVDFNTIRIFPSGNV
 TNILEMRNKDFRHSHPAKHYLATDPMGSAVFLSDTNSRRVKVKTVMVKDLVNS
 EVVAGTGQCLPFDOTRCDGCKATEALTINPRGITVDKFGLIYFVDGTMRVDQNG
 IISTLLGSDNLTARPLSCDSMBEISOVRLEWPTDLAINPMDNSLYVDNNVLOISE
 NHQVIRVAGRMHCQVPIDEFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKN
 RIRQVTSBELSVAGSPGCDCKDANDANCSFGDGDGYAKDAKLNPESLAVCADGEL
 IVADLGNIRIRIFRNKPFNTQNNYELSSPIDQELYLFTSGKHLVTSQSLPTGDLV
 NFTYTGDDDIITHIDNNGMVNRDSTGMPLEWLVPDQVYVWTMTGNSALRSVTTQ
 GHELAMTYHSGSLGATKSNENGWMTTPEYDSFGLTNVTFPTGQVSSFRSDTSSV
 HVQVETSKDQVITTNLSASGAFTVLLQDQVRNSYIAGDSRLULLANGMEVALQT
 EPHLLAGTVNPTVKRNVLPIDNGLMVLWEQRKQARGQVTVFGRRLRVHNRLLS
 LDFRVRTEKTYIDHRRKFTLILYDQAGPSLWSPSSRLNVNVTSPGSHAGIOR
 GIMSERMEYDQAGTSIRIFADGKWSYVILEKSMVLHLSHQROYIFEFKDNRLSSV
 TMPNARQTLTIRSVGYRNIYQPPENASVIOQDFTEDGHLHFLVLTGTRRVLYKY
 GKLSLAETLYDTTKVFTYDGTAGMLKTNLQNEGTCTIRYQIGLPLDRQLFRFT
 EEGMVNARFDYNDNSFRVTSQMAVINETPLIDLYRYDVSQGTGFGKPGVLIYDI
 NOLITTAVMTHKHFDAYGRMKEVQYEIFRSLMTYMTVVDNMGVVKELKGPYAN
 TTRYSEYDADQQTYSINDKLWRYSYDILNGLNHLSPNSARLPLRYDLDRIT
 RLGQVQYKMDDEGLRQGDGVFEYNAGLILKAYNRASWSVRYRVDGLRRVSSKS
 SHSHLOFFVYDTPNTKVTLYNHSSEITSLYYDQHLFAMELSSGDEFFYACDN
 IGTPFLVFSGTGLMIKQILYATGEIYMDTNPNPQIILYVGHGLYDPLTKLVHMGDRD
 YDLVAGRWTSFDHLEWRLSSNSIVPEHLYMFKNNPISNSQDIKCFWTDVNSWLLTF
 GFLQHNVIYGPYKPDTDAMPSELYVHTQMTQWBNDSKSLGVQCEVQKOLKAFVTL
 EREDQLYGSIITSCQAPETKFFASSGSIFGKGVKFAKDGRTVTDIISVANDGRI
 AAILNNAHYLENLHITDGVYDTHYFVKPGESEDAIILGSGRRTELENGVNTQSL
 NTLMSGRTRYDIOQYRVALCLNTRYGTTVEDEKRVLELQRAVQRAWAREQVRI
 REGEGLRAWTDGEKQVILNTRGVQYDGFVFTSVQYPELSDSANNIHPWRSEMR
 R"

ORIGIN

Alignment Scores:

Pred. No.: 7,52e-106 Length: 8585
 Score: 1031.00 Matches: 189
 Percent Similarity: 99.00% Conservative: 10
 Best Local Similarity: 94.03% Mismatches: 2
 Query Match: 95.73% Indels: 0
 DB: 10 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AB025413 (1-8585)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
 Db 7391 ATCCGCTACACGCGCGCTCTATGATCCACTCACAAGCTGTCTCACATGGCCGACGG 7450
 QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGlyLeuTyrLysHisLeu 40
 Db 7451 GATTATGATGTCGTGCTGACGCTGACACAGCCCGACCATGAATCTGGAAAGCCCTG 7510
 QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
 Db 7511 AGTAGCACAGCATCGTGCCTTTTCATCTCTACATGTTTAAAGACACACACCCCATCAGC 7570
 QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrLeuLeuThrPheGly 80
 Db 7571 AACTCTCAGGACATCAAGTCTTCATGACAGATGTCAACAGCTGGCTCTCACCTTGGGA 7630

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 Db 7631 TTCCAGCTGCACAACGCTGATACCTGGCTATCCCAAGCCAGACACAGATGCCATGGAAACCC 7690
 QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
 Db 7691 TCCTACGAGCTCGTATCACACACAGATGAAAACTCAGGAATGGACACAGCAAGTCTATC 7750
 QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 Db 7751 CTCGGGGTACAGTGTGAAGTTCAGAAAGCAACTCAAGGCTTTCCTTACCTTAGAAGCTTT 7810
 QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
 Db 7811 GACCAGCTCTACGGCTCGACCATCACAGCTGCCAACAGGCCCTTGAGACAAGAAGTTT 7870
 QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
 Db 7871 GCCTCCAGTGTTCATCTTTGGCAAGGGGTCAAGTTTGCCTTGAAAGATGTCGAGTG 7930
 QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
 Db 7931 ACCACTGACATCATCAGTGTGGCAATGAGGATGGCGGAGGATCGCAGCATCTTGAAC 7990
 QY 201 His 201
 Db 7991 AAT 7993

RESULT 11
 AF059485
 LOCUS AF059485
 DEFINITION Mus musculus DOC4 (Doc4) mRNA, complete cds.
 ACCESSION AF059485
 VERSION AF059485.1 GI:3170614
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Wang,X.Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P.,
 Zinsner,H. and Ron,D.
 Identification of novel stress-induced genes downstream of chop
 EMBO J. 17 (13), 3619-3630 (1998)
 98315054
 9649432
 REFERENCE 2 (bases 1 to 9722)
 AUTHORS Wang,X.-Z. and Ron,D.
 TITLE Direct Submission
 JOURNAL Submitted (14-APR-1998) Skirball Institute, New York University
 MEDLINE Medical Center, 550 First Ave., New York, NY 10016, USA
 FEATURES
 Location/Qualifiers
 1..9722
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NIH/Swiss"
 /db_xref="taxon:10090"
 /cell_line="NIH-3T3"
 /note="tunicamycin-treated"
 1..9722
 /gene="Doc4"
 /note="downstream of CHOP 4"
 583..9060
 /gene="Doc4"
 /note="similar to Drosophila melanogaster tenm/odx and
 human gamma-heretulin; type II transmembrane protein"
 /codon_start=1
 /product="DOC4"
 /protein_id="RAC31807.1"
 /db_xref="GI:3170615"
 /translation="MDVKRPYRSLTERRDAERRYTSSSADSEEGKPGKQSYSSST
 LKAYDQDARLAYSRVKDMVQEAEEFCRTGTNFTRELGLGEMTPPHGLTYRTDIGL

PHCGYSMGASSADLADTDLSPBHVLRWGRSTGRSSCLSSRANSLTLTDTTEH
NTTGAFLHCSSASSTDEQSPSPPPANESQORLLGNVQAQPPDSDSEEFVN
SFLVSGSAGSLGVAANDPSSLQNHPLRTPPLPHAHPTNQHAAINSLNRGFT
PRNSPAPDTHLSGSPAGSQAQEPHAQDNWVLSKIPVETRNHLKQFLGWTQDN
LIEMDIFASRRDAGVSDGHFFPKPGTSPFLCTSPGLTSTVSPPLPRST
FSRPAFLKPSKCNWKAALSAIILATVILLAVFAMHFLGNHLPQMEGQM
MYEITDASAPVFTDVSILYPSGCTGLETFDRKGAAGKPSLPEDFSIDSGE
DVGRRASQKTPPTFFWRSQVFIHPVHLKFNWSLGAALVGLVGRKGLPSHTDGF
ELDGRLLTQEARSLGSPQRPSPSSHETGTOYLDSDGHLIAFLVANGKSEV
VSPUTTAIESVDCPNCGNDICSGTCHFLGPDGCRASCPVLCNSGQYMGK
RCLCHSGWKAECVNTQCIDVACSHGTICMGTCICNPGYKGESEEVDMQDPTCS
SRGVCVRCHCSVGMGGTCTPRATCDGCSGHTFLPDTCNCDPQSWTGDCSI
EICCAADCGHGVGGCTDLNGHVCVQLGWRGTGCDTSHETSCGKNDGRLD
HCTTEGPGLCNGVCRCTDLNGHVCVQLGWRGTGCDTSHETSCGKNDGRLD
CMDDPCLQCHVNPCLGSPDLFIIOETAPVSOQLNSFIDRILFLVROSTHS
IPGNPDPGHCVRGQVMTSDGTPLVGNVISFINPLFGYITSRQDSFDLVNNG
LSIILPERAPFIQHTLWLPDRFFMETIVMRHENEIPSRDLNSFANFVSP
SPLTSFASCAEKPIVPEIOALOEIIVAGCKMLSYLSSTPYKSVVRISLTHPT
IPFNMKVLHVAEGRLFRKFAAAPDLSYFIWDTVDVNOVGLSEAPVSGVE
YESCPDLILWKKRTAVLOGYEDIASKUGWSLKHHLNLSGILHKGNGNOVSQ
PVTGSIINGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLVGFNFIRLIFS
GNTVILHMSHSPAHKYIATDPMGSAVFLSDTNSRFRVKSSTTVKDLVKNSEVA
GTDOCLPDDTRCDGKATEALTNPGRGITVDFGLIYFVDTMIRRVQNGIIST
LLGNDLTSARPLSCDSVWEISOVRLEWPTDLAINPMONSLVLDNNVLOJSENHOV
RIVAGRPHCQVPGIDQFLLSKVAIHATLESATALAVSHNGVLYIAETDEKINRQ
VTSGETISVAGAPSGCDCKNDANDCFSGDDGYAKADKLNTPSSLAVCAGELVAD
LGNIRIRIRKNNPNTQNMVYELSSPDIQELYFDTSGKHLITDPTGDIYLNFTY
TGDDGITHTDNNGNVNRDSTGMPLMVLVDPQVYVMTGNSLARSVDTSVHVQ
AMTWKSGDYLITLHSGSAGFYLLQDOVNSYIIGADGSLRLLLANGMEVALQTEPHL
ETSXKDDVITLHSGSAGFYLLQDOVNSYIIGADGSLRLLLANGMEVALQTEPHL
LAGTVNTPVGRKNTLPIDNGLNLMVQRKEQAGQVTFGRPLRVHNRLLSLDPL
RVRTKIYDHRKFTLRLYDQAGRFWSFSSRLNGVNTYSPGHLAGIQRGIMS
ERMEYDQIRITSPADGKMSYIILEKSMVHLHLSQYIFEDKNDRLSVTMPN
VARQLETIRSVYRNYPQPEGNASVQIDFEDHLLHLYLGTGRVIRYKGLIS
KLAETLDTKVSFYDETAGMLKTNLQNEGTCTIRYQIGLIDRQIFFTBEEM
VNAREVYDNSFRYTMQAVNETPLDLYRVDYDVGKTFQKFGFYIYDINQII
TYAVMTXKHFDYGRMEKVOYETFRSLNMYMTVOYDNNVKKELKVPYANTYR
SVYEDADGLQTVSINDKFLYSDYDLNGLHLSPGNSARLTPLRYDLDRITRLGD
HQQVFDADITNPTKTHLYSHSSEITSLYDLQGLFAMELSSGDEFIACDNTGTP
VQYKMDSDLRQGRDVEFYNSAGLLIKAYNRSAGSVRYRVDGLGRVSKSSH
LAVFSTGLMIKQILLYTAYGIEVMTDNPNNFIIGYHGLYDPLTKLVHGRDRIYD
AGRWSPDHLKRLSSNIVPFLHMYFNKNPINSQDIKCFMTDVSMLLTFQFL
HNVIQYXPDDTAMPSRYELVHTOMKTOEWDNSKILGVQCEVQKQALFVTLERD
QYLSSTISCOQAPETKFASSGSIFGKGVKALKDGRVTTDIISVANEDGRRIALF
NNAHLENHFTIDGVDTHYFVAPGSGEDLAILGSGRRLNGVNTYVQIINTML
SGRTREYTDIOYRQALCNTRYGTVDDEKVRVLELQARQVRAQWABEQRLRGE
EGLRWATGCKQVILNTRGVQYDGFVTSVEQYPELSDANNHFMROSEGRR
1813. .1881
/gene="Doc4"
/note="putative; transmembrane-region site"

ORIGIN

Alignment Scores: 8.73e-106 Length: 9722
Pred. No.: 1031.00 Matches: 189
Score: 99.00% Conservative: 10
Percent Similarity: 94.03% Mismatches: 2
Best Local Similarity: 95.73% Indels: 0
Query Match: 10 Gaps: 0
DB: 10
US-10-029-020-14_COPY_2400_2600 (1-201) x AF059485 (1-9722)
Qy 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
Db 7948 ATCGGTACACGCGCGCTCTATGATCCACTCAACAGCTTGTCACATGCGCGCAGG 8007
Qy 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 8008 GATTATGATGCTGGCTGGAGCGTGGACAGCCGACGACCATGAACCTCGAAACGCGCTG 8067
Qy 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 8068 AGTAGCAACAGCATCGTGCCTTTTTCATCTACATGTTTAAAGAACACAAACCCCATCAGC 8127

Qy 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 8128 AACTCTCAGGACATCAAGTGCCTTCATCACAGATGTCACAGCTGGCTCCTCACCTTTGGA 8187
Qy 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 8188 TTCCAGCTGCACACGTCGATACCTGGCTATCCCAAGCCGACACAGATGCCATGGAACCC 8247
Qy 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
Db 8248 TCCTACAGAGCTCGTACACACACAGATGAAACTCAGGAATGGGACACAGCAAGTCTATC 8307
Qy 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 8308 CTCGGGTGTACAGTGTGAAGTTCAGAACTCAGAGCTTCCTTACCTTGAACGCTTT 8367
Qy 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 8368 GACCACTCTACGGCTCCACCATCACCAGCTGCCAACAGGCCCTCGACACAAAGATTT 8427
Qy 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 8428 GCCTCCAGTGGTTCATCTTTGGCAAGGGGCTCAAGTTTGCCTTCAAGATGCTCGAGTG 8487
Qy 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
Db 8488 ACCACTGACATCATCAGTGTGGCAATGAGGATGGCGGAGGATCGACGCACTTTGAAC 8547
Qy 201 His 201
Db 8548 AAT 8550
RESULT 12
AB026980
LOCUS
Danio rerio mRNA for ten-m4, complete cds.
AB026980
ACCESSION
AB026980.1 GI:5307784
VERSION
ten-m4.
KEYWORDS
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (sites)
Mieda,M., Kikuchi,Y., Hirate,Y., Aoki,M. and Okamoto,H.
Compartmentalized expression of zebrafish ten-m3 and ten-m4,
homologues of the Drosophila ten(m)/odd Oz gene, in the central
nervous system
Mech. Dev. 87 (1-2), 223-227 (1999)
JOURNAL
MEDLINE
99425191
PUBMED
10495292
REFERENCE
2 (bases 1 to 9264)
Mieda,M.
Direct Submission
Submitted (06-MAY-1999) Michihiro Mieda, Brain Science Institute,
RIKEN, Lab. for Developmental Gene Regulation; 2-1 Hirosawa, Wako-shi,
Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp,
Tel:81-48-467-9713, Fax:81-48-467-9714)
Sequence updated (29-Jun-1999).
COMMENT
FEATURES
Location/Qualifiers
1..9264
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/dev_stage="embryo"
1..9264
/gene="ten-m4"
323..8797
/gene="ten-m4"
/note="similar to Drosophila melanogaster tenm/odx and Mus
musculus Doc4"

```
/codon_start=1
/product="ten-m4"
/protein_id="BAA81893.1"
/db_xref="GI:5307785"
/translation="MEVKERPYRSLTSRRDTRRRTYSSADSDGKINPKSYSSSET
LKAQDQSRSLAYGRVLDVHDEAFSRQDFSLRDMAFDQVPPHMGAYRTEMGL
PHRYSVSVAADATETDIPSPHVAHVRMGSRNLTSGRSLFRANSNLTLDTEH
ENTENGPLHCSASSGDSPIYPPHAAQNSOGRLLGNSQAQGRDSEDEDFGN
SFLVKTGSNNVCTPAATANEQSGSQDNLLNSVPLETRNTAKOTFLELQD
NFIEMDILARRDAYTDGHFLFKPGSTSLYCTTSPGYELTSTSTVYPPRLPN
TFSPAPSLKPKYKHCNMKCAALISLTVFLFLAYFAMHLFNHLPQVOROI
YQLTEDNTGHLPTDLGLPLNGTGLEFPPDRSGDKLDPFPEDSFIMDELIVG
RKVAQLIPGPIFWSQVPIIDHPMYLKNVSLKDALYIGRGLPSPHQFDFVELL
DGRLLSQPLGDPPPPAAQORSLVLTSHDTGCIQYMSGSIWHLAVINDGKRETV
SFLTADSIDDCPSNCFNGDCVSGNCHFPFRGDCSRASCPVLCSNGNGYKOR
CMCHSGSECDVTQCIDITCSGHGTCTVGTICINPSYKGENCEVEFCLDPTCSG
RGVCVRGECHFCVWGQPGCESPRASCMQCGHGSFLADTNTCNCNHNWTHDCSTE
LCAADCGGHI CVAGSCRCDEGMWGTGCEQACHPCRSHEGTCKDGKCEGSPWNGEH
CTTGCPGLNGNRCITLNGNGWYCVQOLGWRGAGCDSMETACSDGKNDGDLTDC
MDPDCLOASCHTSLCVGSPDLDIIOETQISLSSLTQSFYQRIHFLVGRDSTHVI
PVNPPDGIHACTVIRGOVTSDEGPLVGVNISFINKPAIGTITIRQDGSFEDLVNNGV
AIGLFRAPFITEHITLMPGWFVMDTIVMKHEVNDIPSCDLSQSTFPMPTLPA
PLFAAGTCBERGIVPEIOTLOEVRIPGTDMLGYSRTSGYSKSLRITLTHSTI
PFLMKVHLMAVEGRILKWFSAAPNLSDYDFVMDKTDVYSQKYGSLSEAFVSGREY
ESCPDLILWEKRTAVLOQYETTASNLGWSVDKHALNIOGILHKGNGENIPISQP
PVTGSLMNGRRRSICPSCNGLADGNKLLAPVALACGSDGSLVGVDFNVRRLFTTG
NVTSLVLSNPAKHYLATSPVSGWLYLSDTSRKFVKSLYAVKDVAKNLELVAG
TGQCLPYDTRCGDGAKEATLTPRGITVDKYGVFFVDGTMIRRIDQNGLISITL
LGFNDTSLAPLSYMDISQVRLEWPTDLAVSPMNSLYVLNDNNVLIQISNHOVR
IVAGRPMHCQVGLDHLFVSKIAHATLESANALAVSHNGLLYIAESDEKIKNRQV
STNGETSLLAGAPGCDKNDANCDCYGDGAKOAKLAPSLAVSPGELFIADL
GNIRIRVRNKAFLPNLNNVEJSSPIDDELXLFDVNASHVFTQSLITGDIYLFYIS
GEGDLASITDKNKNRYSIRDSITGLPLMLMGPDQTFWFTMGNTNALKSVAAGQSEIA
VMTYHSSGLLAKSDGEMSTFEYDNYGRLNTVPTGRVSSYRTSDSDTSTVROTE
GSKNKDITVTNLSAGTFYTLMDQDQVNSYIIGLDSRLVLVLANGVMSLWTEPHLL
SGTNVPTIKRNTVITDNLGNLEWQRKEQAGQVTVYGRRLRVHNRLLSMDFDR
VTEKDYDDHRTFLIRHYDHAGRPTLAPSLRNGVNTVYSGPGHLAGIQRTWSV
RMEYDQNRITSKI FADGKSWSYTLEKSMVLLISQRYVIFEDKKNRLUSSVTMNV
AROLTETRSIGYRNTYRPPENATVQYSEDGLLQTHQGTGRVRYIKYGLSR
LLEILVDTIRAFSDSAGMLKVLQSEGFACITRYQIGLIDRQIFRFSSEGMV
NARFDYNSFVTSQVINETPLFIDLYRYDVSQKTEQFGKFGVYIYDINQIIT
YEMDQGLQVYNSINDPLWYSYDLNGLHLLSPGNSARLPLRVDIRIRLQDV
TAVDQKTHFDAYKRVEQYEIFRSLYMMVQFONMGRVAKELKVGPIYANTYVA
QVLEDDGFLRQNDPFEVNSAGLLVYKYNKNTIKYRVDGLRRVSRSTQGH
LOFFYADLASGTVMYHNSSEITSLYIDQHLFAMELSSGDEFFVACDNIGPLA
AVFSAGLMIKQLHTAFGEVYLDNSPFLVIGYQGLYEPLTKLVHNRGRDYDVL
GRWTPDHDIRKLNDSNIPVFNLMKNPNLSQETKCYMTDVSMLVTFPGFOLY
NIVPGYKPTVDAMESYELVHTQIKTOEWDSKSVLQVCEVOROLSKFRLERFGQ
IYISADSGCPPTPLHTFATGSLFGKGVKVAIREGRVADIIISLANEDGRRTAAVLD
KASYLQDLHFTTAGLTHYFVKSGLVGDLISLGMVTVGORTLETGVTNVTVSQNMVLG
GRSRRITDIOMQYGTLSLNVYSGSVDEKRVRLVLEARAVATAWAHRRHLRQEE
GSRWTDGERQLLSGRVQGEFGFYIVSDQFPFELTDNINNHFVWQTEMGR"
```

Qy 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7805 AACAGCGCAACATGTGTCCTTCAACCTTACATGTTTAAAGAACAAACACCCCTTGAGC 7864

Qy 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7865 AACAGCCCAAGAACCAAGTGTATACATGACATGTTAAACAGTGGCTGGTACGCTCGT 7924

Qy 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7925 TTCCAGCTATACACGTCATCCCTGGCTACCGAAGCCTGTACAGATGGCATGGAGCT 7984

Qy 101 SerTyrGlnLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
Db 7985 TCGTACGAGCTCGTTCACCCAGATAAAGACTCAGGAGTGGATAGCACCACCACTGTT 8044

Qy 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 8045 CTAGGGGTGCAGTGTGAAGTCCAGAGCAGCTGAAGTCTTTTGTCCGCTTGGAGCGGTT 8104

Qy 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 8105 GGTTCAGATCTACAGTGCAGTGTCTTGGATGCTCTCAACCCCTCTCCACACACTTTT 8164

Qy 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 8165 GCAACAGGAGCTTCTCTCTCGGAAAAGTGTGAAAGTCGCCATCCCGTGAAGCGCGTT 8224

Qy 181 ThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAlaLeuAsn 200
Db 8225 GAACCGGACATCATCAGCTGGCTTAATGAAGATGGCCGAAGGATCGCCGACGCTAGAC 8284

RESULT 13
GGA238613 8118 bp mRNA linear VRT 02-JUN-1999
LOCUS Gallus gallus mRNA for teneurin-1.
DEFINITION Au238613.1 GI:4877312
ACCESSION ten-1 gene; teneurin-1.
VERSION Gallus gallus (chicken)
KEYWORDS Gallus gallus
SOURCE ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1
AUTHORS Minet,A.D., Rubin,B.P., Tucker,R.P., Baumgartner,S. and
Chiquet-Ehrismann,R.
TITLE Teneurin-1, a vertebrate homologue of the Drosophila pair-rule gene
ten-m, is a neuronal protein with a novel type of heparin-binding
domain
J. Cell. Sci. 112 (Pt 12), 2019-2032 (1999)
JOURNAL MEDLINE 99276585
PUBMED 10341219
REFERENCE 2 (bases 1 to 8118)
AUTHORS Chiquet-Ehrismann,R.
TITLE Direct Submission
JOURNAL Submitted (28-APR-1999) Chiquet-Ehrismann R., R-1066.446, Friedrich
Miescher-Institute, Postfach 2543, CH-4002 Basel, SWITZERLAND
FEATURES
Source
1..8118
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/tissue_type="brain, retina"
/dev_stage="E11, E14, adult"
1..8118
/gene="ten-1"
1..8118
/gene="ten-1"
/function="putative role in neuronal development"
/codon_start=1
/product="teneurin-1"
/protein_id="CAB43098.1"

gene
CDS

Qy 1 IleGlyTyrHisGlyCysLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7685 ATAGGCTACAGCGCGCTTATGAGCCCTTACAAAGCTTGTCAATATGCGCGCAGA 7744

Qy 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7745 GATTATGATGTTCTTGACAGCGCGGTGGACCACTCTGACCATGATATCGGAAGCGGCTC 7804

ORIGIN

Alignment Scores:
Pred. No.: 1.87e-80 Length: 9264
Score: 807.00 Matches: 143
Percent Similarity: 87.50% Conservative: 32
Best Local Similarity: 74.50% Mismatches: 25
Query Match: 71.93% Indels: 0
DB: 5 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AB026980 (1-9264)

| | | |
|---|---------------|-----------------------|
| /db_xref="GI:4877313" | repeat_region | /gene="ten-1" |
| /db_xref="SFRSMBL:O9W6V6" | repeat_unit | /note="EGF repeat 8" |
| /translation="MEQMDCKYQPLSKVKVHEVDLTYTSSDESDEGRKQROSYDSRE TLNYSQELRLNYSQRRKNQDSTQMECEPETHILCSGYQDHLGVSBSHVSYLE VGSVDVTEGGASDPHALRMWGMKSHSSCLSRANSALSITDIDHERKSGDEND MGSPDHNOFTFRPLPPPPHACTCTRPPPAADSLQRSMTRTQSPSPAAPTPTS TODSVHLHNSWLNSNIPLERHFLFKHGSGSSAIFSAASQNYPLTNTVYSPPRPL PRTSPFAFTSKPKYRCRCNMKCTALSATITVTLALLAYIAVHFLGLTQLOQPL GOIYENSVKNGKAEDTDTTSPIGCKVSDTEKKVFKQGRAIDGVEIKGAQMOT IPGLFWRFQITTHHVPYLFNLSLAKDSLLGIYGRNNIPPTHQTFDFVKLMDGKQLI KQEPKNSRPPQAPARNLITSLOETGFIYDMQGAHMAFYNDKKVEQVFLVITIAIE VLDDCTNCGECCGHCPCPGFLPDCADKSCPVLCSGNGEYKCHGVCGVCGE GPCDVPQOCDICTCFGHCITCMGVCICVPGVKGEICBEECDLCPMCSGHCVCQGE CHCSAGWGVNCEISPICOEHSCHGFTLLDVLGICSCPEQWTSDCSTELCTLDGCG HGVCSRGICQCBEGVGPCEERTCHSHCAEHCQKCKECSRGWEGDHCTIDGCPG LCYNGRCITDQNGHWCVCQVGMWGSCNVVMEMACGDLNLDGDLCTDPCDDCCQ NNCYASPLCGSPDPLDLTQHSOPFPSPHPPFYDIRIFLIGKESHVIGPISDES NRASVAVGTGVLPGVNVYFISLRVDEYGTISRQDGSFLDVAVGGSVTLVDFDR SPISEKRTLWLSNRFVIVDKVMQRAESDIPSCDVSSFISENPVLSPSLTAFGGS CPERCVTPELOVQVREPIPIPSSEFKLYLSRTPGVKTLLRVILTHTIPSGMTKVH LITAFGRLLQKFPAAANLVTFANKDITYGKYSGLAEAMVSYEYETPCDPLIL WEKRTVILQGFENAGNLGWSINKHVLNPSQGIYHKGNGEMFI SQQPPVLSITMNG NQHRSVCSNCGNALNSLAFAPLTPSPDGSVYIGDFNFRIRIPPSGNSIGILEL RNDTRHSTSPAHKYIADVPSESLSYLTNTRRVYKASLSLIETKDLAKNVDVAGT GGDCLPFDQSHCGGKSEASLNSPRGITDKHGFIYFVDMIRKIDENGMITTII CSNGLSTOPLSCDSGMDITQVLEMPDTLTVNPLONSVLVDNNIVLQISESRVRI IAGRIHCQVPGDHFIVKVAIHSTLESARALAVSHGPIYIRETDERKINLQOVT TNGEISIIAGPDCDIDPCDPSGDCGYAKADKLAPSSLA VSPDITLVADLIG NIRIRAVSRNKAHSLNMYEYIADSPADELYQFTINGHTLHLNLTIRDIYINFTYSG EGDVATITSNGNSVHRRITSGLPVLVVVPGGVYMLTSSNGVLKRVYAQGNLAL MTVPGNTLLATKSDENGTWTVYEYDSGHLTNATPTTGEVFSFHSDEKLTVELDT SNRENMTAFNASTSTIYTLKQDNTONIYRVPDGSRLTVTFASGMEITLNTPHILIA GVVSPTLGKNISLPGHNSNLNEMRORBOTKGNISTERRLRNHRNLLSIDPDIHV TRTKIYDDHRDITRIMDQYRPLPSPISPKYNEVNIYSHSGLYTILQRTWTPEK MEYDPGNIISRTWADCKIYSPLEYSVMLLHSQRRYIFEDQSDILYSVTPMSV RHALQTLMSGYRNIYTPDPSGAFTQDVTGRLLQTLTPYGTGRVLYIKSQSRL SEILYDTQVTFVESSGVKTIHLMHDGFICTIRYRQTGELIGROI FRPSEGLVN ARDPYSNNRVPVSMQAMINETPLPIDLYRVVDSGRTEQGFKEFSVINYDINQVITTT VMKTKLIPSAQGVIEVQYELLKSIAYMTIYDNDMGWMTICDIRVGVDAMI TRFYVE YDRDQLOTVSNDKTQWRYSDYDUNGNNILLSHNSARLTLPLRLDLORITRLGEIOY KMBDGLRORGENEIFRYSNGLLNKAYKSGWTVQVCYDGLGRVACSKSLQHLQ FYADLSNPTRVTHLYNHSSEITSLYDLOQHILIA MELSGSEYVACONDYPLAV PSSRGQVKEITLYTPYGEIYODTNPDPQVIFGHGLYDSLTKLHLGQRDYDIAGR WTTNPHHILKHNAPPOFNLYSPENNVPGVRIODVAKYTTDIGSWLELFGFOLWTR PGFKPELEALETYYELLOQTQEWDPGKTIIGIOCELOQLNRNLTSLDQLPMTPR YSDCKEYGVQKQAPALPSVFGKGIKFAKDGIVTADILGVANEDSRIAAILNNAH YLENLHTIEQRDTHYFKIGLSBEDISLITGGRRIENGUNVTVSQTMSVINGRT RFPADIQLOHCAICFVRYGTVEEKNHLEVARQAVAKWTKQRRLQEGEGEGR AWTDGEKQLLANTGRVQDGYDFVLSVEQYLELSDSANNIHFMRQSEGRP" | | |
| 1525..2304 | repeat_unit | /note="EGF-like" |
| /note="EGF-like" | repeat_unit | /gene="ten-1" |
| 1525..1617 | repeat_unit | /note="EGF repeat 1" |
| /note="EGF repeat 1" | repeat_unit | /gene="ten-1" |
| 1618..1710 | repeat_unit | /note="EGF repeat 2" |
| /note="EGF repeat 2" | repeat_unit | /gene="ten-1" |
| 1711..1812 | repeat_unit | /note="EGF repeat 3" |
| /note="EGF repeat 3" | repeat_unit | /gene="ten-1" |
| 1813..1908 | repeat_unit | /note="EGF repeat 4" |
| /note="EGF repeat 4" | repeat_unit | /gene="ten-1" |
| 1909..2013 | repeat_unit | /note="EGF repeat 5" |
| /note="EGF repeat 5" | repeat_unit | /gene="ten-1" |
| 2014..2106 | repeat_unit | /note="EGF repeat 6" |
| /note="EGF repeat 6" | repeat_unit | /gene="ten-1" |
| 2107..2199 | repeat_unit | /note="EGF repeat 7" |
| /note="EGF repeat 7" | repeat_unit | /note="EGF repeat 8" |
| 2200..2304 | repeat_unit | /note="EGF repeat 9" |
| /note="EGF repeat 9" | repeat_unit | /note="EGF repeat 10" |
| /note="EGF repeat 10" | repeat_unit | /note="EGF repeat 11" |
| /note="EGF repeat 11" | repeat_unit | /note="EGF repeat 12" |
| /note="EGF repeat 12" | repeat_unit | /note="EGF repeat 13" |
| /note="EGF repeat 13" | repeat_unit | /note="EGF repeat 14" |
| /note="EGF repeat 14" | repeat_unit | /note="EGF repeat 15" |
| /note="EGF repeat 15" | repeat_unit | /note="EGF repeat 16" |
| /note="EGF repeat 16" | repeat_unit | /note="EGF repeat 17" |
| /note="EGF repeat 17" | repeat_unit | /note="EGF repeat 18" |
| /note="EGF repeat 18" | repeat_unit | /note="EGF repeat 19" |
| /note="EGF repeat 19" | repeat_unit | /note="EGF repeat 20" |
| /note="EGF repeat 20" | repeat_unit | /note="EGF repeat 21" |
| /note="EGF repeat 21" | repeat_unit | /note="EGF repeat 22" |
| /note="EGF repeat 22" | repeat_unit | /note="EGF repeat 23" |
| /note="EGF repeat 23" | repeat_unit | /note="EGF repeat 24" |

```
repeat_unit 6667..6729
/gene="ten-1"
/note="YD repeat 24"
6730..6801
/gene="ten-1"
/note="YD repeat 25"
6802..6867
/gene="ten-1"
/note="YD repeat 26"

ORIGIN
Alignment Scores:
Pred. No.: 9.85e-65 Length: 8118
Score: 667.50 Matches: 118
Percent Similarity: 77.61% Conservative: 38
Best Local Similarity: 58.71% Mismatches: 44
Query Match: 61.98% Indels: 1
DB: 5 Gaps: 1

US-10-029-020-14_COPY_2400_2600 (1-201) x GGA238613 (1-8118)
QY 1 IleglyThrsHisGlyGlyLeuTyAspProLeuThrsLysLeuValHisMetGlyArg 20
Db 7009 ATTGGTTTCATGAGGCGCTCTATGATCTCTACCAAAATGGTTCATCTGGTCAGCGG 7068
QY 21 AspTyAspValLeuAlaGlyAigTriThrsSerProAspHisGluLeuTrpLysHisLeu 40
Db 7069 GATTATGATGTTATGCTGCTGCGTGACACACCAACCAACCATCATCATGAAACACCTG 7128
QY 41 SerSerSerAsnValMetProPheAsnLeuTyMetPheLysAsnAsnAsnProIleSer 60
Db 7129 AATGCTGTC---CCACAAACCATTCATCTCTACTCATTTGAAAATACTACCCAGTTGGC 7185
QY 61 AsnSerGlnAspIleLysCysPheMetThrsAspValAsnSerTrpLeuLeuThrsPheGly 80
Db 7186 AGGATCAAGATGTGCTAAGATATACACAGACATTCGAAAGTTGGCTAGAGCTATTITGT 7245
QY 81 PheGlnLeuHisAsnValIleProGlyTyTrpProLysProAspMetAspAlaMetGluPro 100
Db 7246 TTCCAGTTGCACATGTACTACTCTCAACAAAACCCAGAGTGGGATCTCGAAGACTATTC 7305
QY 101 SerTyGluLeuIleHisThrsGlnMetLysThrsGlnGluTrpAspAsnSerLysSerIle 120
Db 7306 ACATATGAATCTTACAGCTTCAACAAAACCCAGAGTGGGATCTCGAAGACTATTC 7365
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrsLeuGluArgPhe 140
Db 7366 CTGGTATTACAGTGTAGCTACAGAACCACTCCGAAACTTATATCTTGGATCAACTT 7425
QY 141 AspGlnLeuTyGlySerThrIleThrsSerCysGlnGlnAlaProLysThrsLysPhe 160
Db 7426 CCAATGACCCCGAGGTATAGTGTGCAAGTGTATGAGGAGTGAAGCAACCGAGGTTT 7485
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 7486 CGAGCTATTCTTCATGTAATTCGAAAGGCATCAAAATTTGCTATCAAGATGCGATCGTC 7545
QY 181 ThrThrAspIleSerValAlaAsnGluAspGlyArgValAlaAlaLeuLeuAsn 200
Db 7546 ACAGCGGACATTATTGTTGCTGCTAATGAGGACAGCCGCGCATCGCTCCCTACTCAAC 7605
QY 201 His 201
Db 7606 AAT 7608

RESULT 14
AB025410
LOCUS
DEFINITION Mus musculus mRNA for Ten-m1, complete cds.
ACCESSION AB025410
VERSION AB025410.1 GI:4760775
KEYWORDS Ten-m1.
SOURCE Mus musculus (house mouse)
```

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (sites)

AUTHORS

Ohashi, T., Zhou, X., Feng, K., Richter, B., Moergelin, M., Perez, M.T.,
Su, W., Chiquet-Ehrismann, R., Rauch, U. and Faessler, R.

TITLE

Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues

JOURNAL

J. Cell Biol. (1999) In press

REFERENCE

2 (bases 1 to 8373)

AUTHORS

Ohashi, T.

TITLE

Direct Submission

JOURNAL

Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University

Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,

Okayama, Okayama 700-8252, Japan

(E-mail: ohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128,

Fax: +81-86-222-7766)

FEATURES

Location/Qualifiers

1..8373

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Balb/c"

/db_xref="taxon:10090"

/dev_stage="adult"

/tissue_lib="brain"

1..8373

/genes="ten-m1"

96..8291

/genes="ten-m1"

/codon_start=1

/product="Ten-m1"

/protein_id="BAA77396.1"

/db_xref="GI:4760776"

/translation="MSQTDCKVQPLKVKHMDLAYTSSDSESDGKRPQSFNPRE
TLHEYNOELRNYSQSRKQVEKSTQLEFCETPTLCSGHTLDMHVSRRGQLE
MGSVDVTEGASPDHALRMIRGMKSHSSCLSSRANSALSTLTDHERKDGENG
RFPSPCCDMEAPADSAQMQSPHNOFTFRLPPPPPPHACTKPKPTVDSIQ
RMTQSPSPAPAPTSTQDSVHLHNSVLNSNPLETRHFLFKHGSSSAIFGAA
SNQYPLNTVYSPRPLPRSTSRPAFTENKPYRCNWKCTALSATATVTLALL
AVVIAHLGLTWLQVQGIYANGISNGNPGTESMDTYSPIGGVSKSEKVFQK
GRAIDTGEVDIGAQMWTIPGLFWRPQITIHPIYIKFNISLAKSLIGIYGRNIP
PHTQDFVKLMDGKQLVQDSKSDDIQHSFNLNLTSLQETGFTIYMDQGWLYAF
YNDKMEQVFLVITTAIEIMDDSCNCGNCGECISGHCHFCFGLDPCARDQSPVLC
GNGEYKGVCHVCRNGKGECDVPEQICDPTCFHGTICMGVCIYCPYKGEICBE
EDCLDPMCSHGIYKVGECSTGWCNCTPLPICQSCSGHGFLLDITGVCSQD
KWTGSCSTELCTMECSHGVSRGICQCEGWSGPTCERSCHSCHAEGQCKDKC
ECSWEGDHCTIAHVLDAVROGCPGLCFGNRCITLDQNGHCVQVSGWSTGCLVM
EMLCGNLNDGDLTCDVDDCCQSNCTVSPCLQSPDLDLIQSQSPFLFSQTSR
LFDRIKLGKDSHTHVQDIDISFDSRRACVIRGOVAVDGTPLGVNVSFLHSDYG
FTISQDGSFDLVAIGGISVLIIFDRSPFLSEKRTLWLPNQIIVVEKVMQRIADA
PSCDISNFIPLPPLTSPFGSGCPERTIVPELVQVQBEIPFSSFVRLSYLSS
RTPGYKTLRIILLTSTIPVGMKVHLTVSVRGLTKWFPRAINLVYTFANKTDIY
GQKWMGLAALVSVGVYEMCEPFIIMWEQRTVVLQGFENDASNLGWSLKHFIHFQ
SGIHLKNGENMFISQPPVIATIMNGHRSVACTNCPNPAHNNKLPAPALASGPD
TKVYKLSLKVETKOLSKNFVAGTQCLPFDQSGDGGKASEASLNSPRGITVD
RHGTFYVDGTMIRRIDENAVITTVISGNTSTQPLSCDGMIDITQVLEWPTDLAV
NPMDSLVLDNNIYVLOISSENRVRIIAGRPIHCOVPGIDHFLVSKVAHLSBARA
LSVSHGILFTAEITDERKVNRIQVTTNGEISIIAGAPTDCCDKIDPNDCFSGDDGY
AKDAKMAPSLAVSPDGLTVADNGVIRITISKNQHLNMLYELIASPADQELVQ
FTVNGHLNMLITRDVYNYFTNAEGLDGAITSNGNSVHIRRDAGMWLWVVP
GOVWLTISNGVLRVSAQYNALMTYPGNTGLLTKSNENGWTVYVDPGEHLT
NATPFTGEVSHFSDLEKLTVALDNRNENFLMNTATSTIYILKQENTSTYRV
SPDGLRVTAFASGMEINLSSEPHIYLAGAVNPLTGCNLSLPGENANILEWQREVN
KYNVSAPERRLAHNRNLSIDFDMHTGTGIYDHRKFTLILYDQGRPLWSFQV
RYNEVITYSPSGLVTFQRTGWNEKMYDQSGKIISIRYADGKIYSLYLEKSWML
LHSQRYIFEQDQSLSTWPSNRVSLQMLSVGYRNIYTPDSTSTFIQDYSR
DGRLLQHLQDGRRLVYKTKQARLSILYDTQVTLTYESSGVITKIDHMGFI
CTIRYQTLQGLIGRQIFERSEGLVNRFDYSNNFRVTSQVAINTEPLHLYRVY
DVSGRTEQFGKFSYVNDLQVITTVMKHTKIENANGVLEVOYELIKALAYWMTIQ
YDNMGRMVICDIRYGVADANITRYFVENDQQLQTVSVNDKIQRYSYDLNKNILLS
HGNSARLTPLYRDLRDLRITRIGETIQYKMDGDFLRQGNDFIRVTHLYNHTSABITSLYDQCG
GWTQYIYDGGRRVASKSLSLQQLQFFADLANPIRVTHLYNHTSABITSLYDQCG

HLIAMELSGEEVYVACDNNMGTPFLAVSFSGOVIKELLYTPGYDVIHDTVPDEVIIG
FHGLYDFLTKLVHLGQDIDVWAGRWITENHHIWKQNLNLPFPNLYSENNYVCK
IQVAKYTTDITGWLFLGQHNVLPGFPKPELENMELIYLLQKTKQEWDPGKM
ILGIQCLQKQNRNFIISLDQPMPTQYNEGRCLGGKOPFAAVQSPVFGGKFAIKE
YGTADITIGVANDSRLAALNNNAHLENLHFTIEGRDTHYFKLGLSDEEDVLIGN
TGGRRILEGVNTVTSQMTSVLNGRTRRFADIOLOHGCALCFNIRYGTVEEKNHVL
MAKQRAVAQAWTQERLQGEGBRTRWTEGEKQQLLGTGRVQGDYGVFLSVBQYLE
LSDSANNHFMROSEIGRR"

ORIGIN

Alignment Scores: 3.3e-61 Length: 8373
Pred. NO.: 636.50 Matches: 116
Score: 75.37% Conservative: 37
Percent Similarity: 75.37% Mismatches: 45
Best Local Similarity: 57.14% Indels: 5
Query Match: 59.10% Gaps: 2
DB: 10

US-10-029-020-14_COPY_2400_2600 (1-201) x AB025410 (1-8373)

Qy 1 IleglyTyHisGlyLeuTyAspProLeuThrLysLeuValHisMetGlyArg 20
Db 7182 ATTGGTTTCATGGAGACTATCATTTCTTACTAAACTAGTCATCAGGCAAGA 7241
Qy 21 AspTyAspValLeuAlaGlyArgThrSerProAspHisGluLeuTriLysHisLeu 40
Db 7242 GATTATGATGTGTGTGGCAGATGGACAGCCGTATCATCATCATATGGAACAGTTG 7301
Qy 41 SerSerAsnValMet-----ProPheAsnLeuTyMetPheLysAsnAsnAsnPro 58
Db 7302 -----AACCTTCCTTAACCACTCACTCTCTCTTTGAGATAAATACCA 7352
Qy 59 IleSerAsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThr 78
Db 7353 GTTGCAAAATTCAGATGTTGCAAGTATACACAGATATCGGACCTGGCTGAGCTC 7412
Qy 79 PheGlyPheGlnLeuHisAsnValIleProGlyTyTrpLeuProAspMetAspAlaMet 98
Db 7413 TTGTGTTTCCAGTTACATAACACTACTACTGATTTTCCAAACCAAGAACTTGAACATG 7472
Qy 99 GluProSerTyTrpGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLys 118
Db 7473 GAATTAACCTATGAGCTTCTTCACTGACAGAAACCAAGAGTGGCATCTCGAAG 7532
Qy 119 SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGlu 138
Db 7533 ATGATCTCGGCATTGAGTGTGAGCTCCAGAAACCAAGAAATTCATTTCTTGGAT 7592
Qy 139 ArgPheAspGlnLeuTyTrpGlySerThrIleThrSerCysGlnGlnAlaProLysThrLys 158
Db 7593 CAGCTTCCATGACTCCTCAGTACAAATGAGGAGGTGCTTGAAGAGGAAACAGCCG 7652
Qy 159 LysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGly 178
Db 7653 AGCTTTCCTGCTGCTCCTTCTGCTTTCGCAAGGATCAAAATTTGCCATCAAGAGGCG 7712
Qy 179 ArgValThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAlaIle 198
Db 7713 ATAGTCACAGCTGATATTATAGGATAGCAATGAAGATAGCAGCGCTCTGCTGCCATT 7772
Qy 199 LeuAsnHis 201
Db 7773 CTCAACAAT 7781

RESULT 15
AF100772
LOCUS AF100772 8297 bp mRNA linear PRI 30-NOV-1999
DEFINITION Homo sapiens tenascin-M1 (TNM1) mRNA, complete cds.
ACCESSION AF100772
VERSION AF100772.1 GI:6165844
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 8297)
Brandau,O., Schuster,V., Weiss,M., Hellebrand,H., Fink,P.M.,
Kreczy,A., Friedrich,W., Strahm,B., Niemeyer,C., Belohradsky,B.H.
and Meindl,A.
Epstein-Barr virus-negative boys with non-Hodgkin lymphoma are
mutated in the SH2D1A gene, as are patients with X-linked
lymphoproliferative disease (XLP)
Hum. Mol. Genet. 8 (13), 2407-2413 (1999)
20025749
MEDLINE 10556288
PUBMED
REFERENCE 2 (bases 1 to 8297)
Brandau,O., Ohashi,T., Faessler,R. and Meindl,A.
Direct Submission
TITLE
Submitted (22-OCT-1998) Medizinische Genetik, LMU-Muenchen,
Goethestr. 29, Muenchen 80336, Germany
JOURNAL
Location/Qualifiers
FEATURES
1..8297
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xq24-q25"
1..8297
/gene="TNM1"
65..8242
/gene="TNM1"
/function="putative receptor molecule"
/codon_start=1
/product="tenascin-M1"
/protein_id="AAF04723.1"
/db_xref="GI:6165845"
/translation="MEOTDCKPYQLPKVHEMDLAYTSSSESDGRKPRQSYNSRE
TLHEYNQRLMNYVSQRKREKSTQEMFECTSHLSCGYQTDMSHVSHGYLE
MGSDVITEGNAQSDHALRMKIRGKSEHSSCLSSANSLSLTDTHERSKDSBENG
PKFSPVCDMEAAGSTQDVQSSPHNQTFRLPPPPHACTCARPPPAADSLQR
RSMTRSPQSPAAAPPTSTQSVHLHNSWLNINPLETRHSLSFKHSGSSAISAA
SONYPLTNTVYPPRPLPRSTFSPAFKPYRCNCKMKTALSAITVTLALL
AVIAVHLFGTLWQVLEGYANGVSKGRSTMDTTPSIGKVSDEKSEKVFQ
KGRATGCTVDIGAOMOTIIPGLFWRQITIHPIYLFNLSLAKSLGTYGRNI
PPTHQFQVKLMDQKLVQDSKSDTQSHSPNLIILTSLOETGFIEYMDQGPWLA
FYNDGKMEQFVLITALEIMDDCTNCNGEGCSLCHGCPGPGPDPCARDSCPVL
CEDCLDPMCSNHGICVKEGCHCTGCGVNCETPLPVOECQSGHGTFLDDAGVSCD
PWTGSDCSTELCTMECSHGVCSCGICQCEEGWGTPTCEERSCHSTEHQCKDGK
CECSFGWGDHCTIAHYLDVADGCGPLCFGNRCRTLQDGMHVCQVGSQTCNVV
MEMLCGDLNDGDLTDCVDPDCCQSNLYSLPLCOGSPDLDLIIQSQSTLFSHSDY
RLFYDRIKFLDGKSTHVIPEVSFDSRRACVIRQVVALDCTPLGVNVNPLHSDY
GFTISROGSDFLVAIGSILIFDRSFLPEKRTLMPLWNOFIVVEKVTQGVYSD
PSCDINSPIPNPILSPPLTSFGSCPTERTIVPELQVQVEEIPISPSFVRLSL
SPTPGVKTLLRILLTHSTIPVGMIVHLTVAVEGRLLTKWFFAALNVLVTFANKTDI
YQKVGAAELVSVGYEYETCPDFILWEQRTVLQGFEMGNLGDMSLNKHHILNP
QSGIHLKNGENMFISQOPPVISTLMNGHQRVACTNCGPAHNNKLPAPVALAGP
DGSVVDGFNRRIFPSGNSVLSLELSTSPAKHYLANDPVSESLYSLDTRTKVYK
LVFDGMIRKIDENAVITVIGSLTSTQPLSCDGMITQVRLHEPTDLVNPMDNS
LVLDNNVLIQISENNRVRIIAGRPHTQCPGIDHELVSKVAHSTLESARALSVSHS
GLLTAETDERKVNRIQQVTTNGEYIIAGAPTDCDKDPCDCPSGSGGAKAKM
KAPSLAVSPDGTLYVADLGNVRIIRTI SRNOAHNDMNIYEASPADQELYQFTVNGT
HLHNLITRDVYNYFTNSSEGDLGAI TSNGNSVHIRRDAGMPLWLVPQGVYWL
TISSGVLKRVSAQGNPALMTYPNGTGLLATKSNENGTIVYVDPGHLTWTPT
GEVSSFSLDKLTKVELDTSNRENVLMSTNLATSTIYLKOENTQSTYRNPDSGL
FEERLAHNRNLLSIDFHTIKTYIDDRKFTLLIYDQTRPILWGFVSRYEVN
RVTFAAGMBIGLSEPHILAGVNPGLCNISLPGENANLIEWQRKEQNGKNSA
ITYPSGLVTFIQRTWPNKMEYDQSGKIISRTWADGKIWSYTLSEKSMLLHSQR
YIFEDQSCLLSVTPMSVRSLVLTQVTLTVESSGVKTIHLMDGDFICTYRVR
TULHGTRVLYKATKQARLSEVLVLTQVTLTVESSGVKTIHLMDGDFICTYRVR
EQGFLGRQIFRSEEGVNFARFDYSNNFRVTSMAVINETPLPIDLYVYVDSGT
EQGKFEVINYDNLQVITVVMKHTIFSAAGQVIEVQYELKAIATWMTIQDNGVR
HGNMTCRVGVGDANITRYFYEADGQLTQVSVNDQWRYSDYDLNGILLSHGKSAR
LPLRYDLDRITRLGEIOYKMDGDLQRQNDIPEYNSGILLQKAYNSAGMTVOY
YIDGLGRRVASKSLQGLHQQFVVDATANPFRVTHLYNHTSSEITSLYDQLQGLHAME

1SSGEYVACDNTGTPLAVFSSRGQVIKEILYTPYGDYHDTYDPDFQVILGHGGLY
 DFLTKLVHIGQDYDVAHWHTTAVHIIWKQNLPLKPNLYSPENNYVCKIQYAK
 YTTDIRSWLELFGFOLHNVLPFGPKPELENLELTVELLRQLTKTQEWDPGKITLGIQC
 ELQKOLRNFI SLDQLPMTPRYNDGRCLGGKQPRFAAVPSVFGKIFAIKDGIVTAD
 IIGVANESRLIAALINNAHLYLENHFTIEGRDTHYFIKLGSLDEEDLVLIIGNTGRRRI
 LENGVNVSQMTSLNLRTRRFADIQI QHGALCFNIRYGTTVBEEKNHVLEIARORA
 VAQAWTKQRRLOEGEGEIRAWTEGEKQQLLSTGRVQGYDGYVLSVEQYLELSDSAN
 NIHFMRQSEIGRR"

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 1.56e-60 | Length: | 8297 |
| Score: | 630.50 | Matches: | 116 |
| Percent Similarity: | 74.88% | Conservative: | 36 |
| Best Local Similarity: | 57.14% | Mismatches: | 46 |
| Query Match: | 58.54% | Indels: | 5 |
| DB: | 9 | Gaps: | 2 |

US-10-029-020-14_COPY_2400_2600 (1-201) x AF100772 (1-8297)

| | | | |
|----|------|--|------|
| QY | 1 | IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg | 20 |
| Db | 7133 | ATTGGTTTTCATGAGAGACTATGATTTCTTACTAAATTAGTGACCTGGGCAAGG | 7192 |
| QY | 21 | AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu | 40 |
| Db | 7193 | GATTATGATGTTGTGTGGCAGATGGACACGGCCTATCATCATATGGAACAGTTG | 7252 |
| QY | 41 | SerSerSerAsnValMet-----ProPheAsnLeuTyrMetPheLysAsnAsnPro | 58 |
| Db | 7253 | -----AACCTCCTTCCTTAAACCACTTCAACCTCTACTCTTGAATAATACTACCCA | 7303 |
| QY | 59 | IleSerAsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrLeuThr | 78 |
| Db | 7304 | GTTCGCAAAATTCAGATGTGCAAGTATACACACATCATCAGAGTTGGTGGAGCTA | 7363 |
| QY | 79 | PheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMet | 98 |
| Db | 7364 | TTTGGTTTCCAAATTACAAATGACTACTCTGGATTCCCAAACTGAAATAGAAAATTTA | 7423 |
| QY | 99 | GluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLys | 118 |
| Db | 7424 | GAATTAACCTTACGAGCTTCTACGGCTTCACAGCAAAAACCTCAGAGTGGGATCCTGGAAAG | 7483 |
| QY | 119 | SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGlu | 138 |
| Db | 7484 | ACTATCTGGGCATTGAGTGTGACTCCAGAAACAGCTCAGGAATTCATTTCTCTGGAC | 7543 |
| QY | 139 | ArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLys | 158 |
| Db | 7544 | CAACTACCTTACTCCCGATACAAATGATGGACGGTGCCTTGAGGAGGGAAGCAACCA | 7603 |
| QY | 159 | LysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGly | 178 |
| Db | 7604 | AGTTTGTCTGCTGTCCTCTGTGTTTGGAAAGGTATAAAATTTGCCATCAAGGATGC | 7663 |
| QY | 179 | ArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArGValAlaIle | 198 |
| Db | 7664 | ATAGTAACAGCTGATATATAGAGTAGTACCAATGAAGATGACAGCGCGCTTGCTGCCATT | 7723 |
| QY | 199 | LeuAsnHis | 201 |
| Db | 7724 | CTCAATAAT | 7732 |

Search completed: August 14, 2004, 12:03:02
 Job time : 3600.05 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 1344.2 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_2650_2725
Perfect score: 391
Sequence: 1 SQINTVLNGRRTRYTDIQLQ.....EQQLRLEGEGGLRAWTEGEK 76

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgm2_1/USPTO.spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB-GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODS=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -WAXLEN=200000000
-USRR=US10029020 @CGN 1.1 19065 @runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGPOP=6
-FCGPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.in.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.nam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match % | Query Length | DB ID | Description |
|------------|-------|---------|--------------|-------------|--------------------|
| | | | | | |
| 1 | 391 | 100.0 | 3320 | 9 AK127101 | AK127101 Homo sapi |
| 2 | 391 | 100.0 | 3486 | 9 HSM806114 | BX537983 Homo sapi |
| 3 | 391 | 100.0 | 5269 | 9 HSM800629 | AL080120 Homo sapi |
| 4 | 391 | 100.0 | 8354 | 6 AX556500 | AX556500 Sequence |
| 5 | 391 | 100.0 | 8438 | 6 AX675551 | AX675551 Sequence |
| 6 | 391 | 100.0 | 8624 | 9 AB037723 | AB037723 Homo sapi |
| 7 | 391 | 100.0 | 8645 | 6 AX600210 | AX600210 Sequence |
| 8 | 391 | 100.0 | 8993 | 9 HSM806812 | BX640737 Homo sapi |
| 9 | 391 | 100.0 | 68970 | 2 AC090378 | AC090378 Homo sapi |
| 10 | 391 | 100.0 | 170156 | 2 AP002412 | AP002412 Homo sapi |
| 11 | 391 | 100.0 | 186084 | 9 AP002768 | AP002768 Homo sapi |
| 12 | 386 | 98.7 | 2685 | 9 AK056531 | AK056531 Homo sapi |
| 13 | 371 | 94.9 | 834 | 5 GGA293019 | AJ293019 Gallus ga |
| 14 | 368 | 94.1 | 208760 | 2 AC074048 | AC074048 Mus muscu |
| 15 | 368 | 94.1 | 226256 | 10 AC073599 | AC073599 Mus muscu |
| 16 | 365 | 93.4 | 5583 | 10 AK122490 | AK122490 Mus muscu |
| 17 | 365 | 93.4 | 8585 | 10 AB025413 | AB025413 Mus muscu |
| 18 | 365 | 93.4 | 9722 | 10 AF059485 | AF059485 Mus muscu |
| 19 | 364 | 93.1 | 231970 | 2 AC133064 | AC133064 Rattus no |
| 20 | 364 | 93.1 | 257693 | 2 AC103323 | AC103323 Rattus no |
| 21 | 359 | 91.8 | 68970 | 2 AC090378 | AC090378 Homo sapi |
| 22 | 311 | 79.5 | 134245 | 9 HS1052M9 | AL022718 Human DNA |
| 23 | 308 | 78.8 | 8297 | 9 AF100772 | AF100772 Homo sapi |
| 24 | 306 | 78.3 | 8118 | 5 GGA238613 | AJ238613 Gallus ga |
| 25 | 304 | 77.7 | 8816 | 5 AB026979 | AB026979 Danio rer |
| 26 | 304 | 77.7 | 188655 | 2 BX324137 | BX324137 Danio rer |
| 27 | 303 | 77.5 | 8373 | 10 AB025410 | AB025410 Mus muscu |
| 28 | 303 | 77.5 | 134334 | 10 AL163512 | AL163512 Mouse DNA |
| 29 | 303 | 77.5 | 165860 | 10 AL831716 | AL831716 Mouse DNA |
| 30 | 303 | 77.5 | 172465 | 2 BX324192 | BX324192 Mus muscu |
| 31 | 303 | 77.5 | 250093 | 2 AC114058 | AC114058 Rattus no |
| 32 | 296 | 75.7 | 2157 | 6 AX876360 | AX876360 Sequence |
| 33 | 296 | 75.7 | 2157 | 6 BD156088 | BD156088 Primer fo |
| 34 | 296 | 75.7 | 2157 | 9 AK027473 | AK027473 Homo sapi |
| 35 | 296 | 75.7 | 3270 | 6 AX877449 | AX877449 Sequence |
| 36 | 296 | 75.7 | 3270 | 6 BD156663 | BD156663 Primer fo |
| 37 | 296 | 75.7 | 3270 | 9 AK001748 | AK001748 Homo sapi |
| 38 | 296 | 75.7 | 3614 | 6 AX876525 | AX876525 Sequence |
| 39 | 296 | 75.7 | 3614 | 6 BD156175 | BD156175 Primer fo |
| 40 | 296 | 75.7 | 3614 | 9 AK001336 | AK001336 Homo sapi |
| 41 | 296 | 75.7 | 5804 | 10 AK122513 | AK122513 Mus muscu |
| 42 | 296 | 75.7 | 7816 | 10 AF195418 | AF195418 Mus muscu |
| 43 | 296 | 75.7 | 8473 | 6 AX662357 | AX662357 Sequence |
| 44 | 296 | 75.7 | 8487 | 6 AX662359 | AX662359 Sequence |
| 45 | 296 | 75.7 | 8645 | 6 AX662355 | AX662355 Sequence |

ALIGNMENTS

RESULT 1

```

AK127101
LOCUS Homo sapiens cDNA FLJ45158 fis, clone BRAWH3043034, highly similar
DEFINITION to Mus musculus neuregulin 1 (Nrg1).
ACCESSION AK127101
VERSION AK127101.1 GI:34533862
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J.,
Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K.,
Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H.,
Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K.,
Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3320)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUN-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Location/Qualifiers
1. .3320
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRAWH3043034"
/tissue_type="brain"
/clone_lib="BRAWH3"
/notes="cloning vector: pME18SFL3"
ORIGIN
Alignment Scores:
Pred. No.: 2,75e-33 Length: 3320
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_2650_2725 (1-76) x AK127101 (1-3320)
Qy 1 SerGlnLeuAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 2706 TCCAGATCAACACAGTACTTAATGCGGAGGACTAGAGCTACACAGACATCCAGTCCAG 2765
Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluysAlaArg 40
Db 2766 TACGGGGACACGTGCTTGAACACACGCTACGGGACACGTTGGATGAGGAGGACGCG 2825
Qy 41 ValLeuGluLeuAlaArgGlnAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db 2826 GTCCTGGAGCTGGCCCGGCGAGAGCGCGTGGCCCAAGCGTGGCCCGGCGAGCAGAGA 2885
Qy 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 2886 CTGCGGGAAGGGAGGAAGGCTGCGGCTCGGACAGAGGGGGAGAAG 2933
RESULT 2

```

```

HSM806114
LOCUS Homo sapiens mRNA; cDNA DKFZp686D0412 (from clone DKFZp686D0412).
DEFINITION BX537983
ACCESSION BX537983.1 GI:31874053
VERSION BX537983.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3486)
Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
Direct Submission
Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686D0412) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clones@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1. .3486
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="11q14.1"
/clone="DKFZp686D0412"
/tissue_type="human cervix"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
DH10B; sites SfilA + SfilB"
/dev_stage="adult"
1. .3486
/gene="DKFZp686D0412"
<!--.3217
/gene="DKFZp686D0412"
/notes="hypothetical protein, N-terminus truncated"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAD97943.1"
/db_xref="GI:31874054"
/translation="SFGRLTNVTFPTGOVSSFRSDTSSVHVQVETSSKDDVTITNL
SASGAFYLLQDQVRNSYIGADGSLRLILLANGMEVALQTEPHLLAGTVNPTVKRNV
TLFDNLGLNLEVRQKQEQGVTFGRKLRVNRNLLSLGFDRTVTRTKIYDDHKK
FTLRILYDQAGRPFLMSPSRLNGVNTYSPGGYIAGIQIGIMSERWEYDQAGRITSR
IFADGKTWSTYILEKSMVLLHSGQRQYIFEDKNDRLSSVTMPNVARQTLTIRSVGI
VRNIYOPPEGNASVIOFTEDGHLHTFYLGTRGVYKYGKLSKLAETFDYTKVSF
TYBETAGMLKTNLQNEGFTCTIRQIGPLIDRQIFRTEGVMNARFQYNDNSFR
VTSQAVINETPLIDLYRYDDYSGKTEQFGKGLVYDINQIITTAVMTHTKHFDAY
GRKEVQYEIRFSLMYWNTVQYDNGMRVKKELKVGYPANTYRSYEVADGQLOTYS
INDPLWRSYDGLNGLHLLSPGNSARLTPLRIDRITRLGRLSSKSSHSHLOFFVADLTNPK
GGDIFEYNSAGLLIKAYNRAGSWRYRVDGLGRVRSKSSSHSHLOFFVADLTNPK
VTHLYNHSSEITSLYDLOQHLFAMELSSGDEFYIACDNIPTPLAVFTSGTGMKQI
LYTAGGRIYMDTFNPFQI11GYHGLVDPLTKLVHMRDRYDVLAGEWTSPDHLWGH
LSSNVMPNLYMFKNPNPIINSODIKCEMTDVNSLLTFGFLHNVIPGVPKFMDMA
MEPSYELIHTQMTKQEMDNSKSLGVQCEVQKQLKAEVLEREDOLYSTITSCQAP
KTKFPASSGVSFGVKFALDKQRTVDIISVANDGRRVNAIILNHAHYLENLHFTID
GVDPHYFVFGPSPGDLAILGLSGRRLENGVNTVQINVLNGTRRTYTDIQLQY
GALCLNTRYGTTLDEKARVLELQARVQAWAREQORLREBEGLEGLRAWTEGEGKQV
LSTGRVQGYDGFVVISVEQYPELSDSANNIHFMRQSEMGR"
3486
/gene="DKFZp686D0412"
polyA_site
ORIGIN
Alignment Scores:
Pred. No.: 2,9e-33 Length: 3486
Score: 391.00 Matches: 76

```

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x HSM806114 (1-3486)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
 DB 2855 TCCAGATCAACACAGTACTTAATGCGAGGACTAGACGCTACACAGATCCAGCTCCAG 2914
 QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
 DB 2915 TACGGGGCACTGTGCTTGAACACAGCTACGGGCAACGTTGGATGAGGAGGACGCG 2974
 QY 41 ValLeuGluLeuAlaArgGlnAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
 DB 2975 GTCCTGGAGCTGGCCGCGAGAGCCGTCGCCAAGCGTGGCCCGGAGCAGAGAGA 3034
 QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
 DB 3035 CTGGGGGAAGGGAGGAGGCTGCGGCGCTGGACAGAGGGGAGAG 3082

RESULT 3
 HSM800629 5269 bp mRNA linear PRI 18-FEB-2000
 LOCUS Homo sapiens mRNA; cDNA DKFp56400423 (from clone DKFp56400423);
 partial cds.
 ACCESSION AL080120
 VERSION AL080120.1 GI:5262552
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 5269)
 Wambutt,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
 Direct Submission
 Submitted (16-JUN-1999) MIPS, Am Klopferspitz 18a, D-82152
 Martinsried, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 This clone (DKFp56400423) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES
 source
 1..5269
 /location=Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFp56400423"
 /tissue_type="brain"
 /clone_lib="564 (synonym: hfbr2). Vector pAMP1; host
 XL-2blue; sites NotI + SalI"
 /dev_stage="fetal"
 1..5269
 /gene="DKFp56400423"
 <1..556
 /gene="DKFp56400423"
 /note="strong similarity to C-terminus of mouse DOC4"
 /codon_start=2
 /product="hypothetical protein"
 /protein_id="CAB45719.1"
 /db_xref="GI:5262553"
 /db_xref="SPTREMBL:Q9Y4S2"
 /translation="GANEDGRVVAALTDHAHVLNLFHTIDGVDTHYFKPQPSGDLL
 AILGSGGRTELENGVNVTVSQINVLNTRTRYTDIQLQYALCLNTRYGTLDEEK
 ARVLEARQAVRQAWREGEGLRAWTEGEKQKVLSLSTGRVQGVDFYFVVIS
 EQYPELSDSANNLHFMRQSEMGRR"

gene
 CDS

polyA_signal 5226..5231
 /gene="DKFp56400423"
 5251
 /gene="DKFp56400423"

polyA_site

ORIGIN

Alignment Scores:
 Pred. No.: 4,55e-33 Length: 5269
 Score: 391.00 Matches: 76
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x HSM800629 (1-5269)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
 DB 194 TCCAGATCAACACAGTACTTAATGCGAGGACTAGACGCTACACAGATCCAGCTCCAG 253
 QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
 DB 254 TACGGGGCACTGTGCTTGAACACAGCTACGGGCAACGTTGGATGAGGAGGACGCG 313
 QY 41 ValLeuGluLeuAlaArgGlnAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
 DB 314 GTCCTGGAGCTGGCCGCGAGAGCCGTCGCCAAGCGTGGCCCGGAGCAGAGAGA 373
 QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
 DB 374 CTGGGGGAAGGGAGGAGGCTGCGGCGCTGGACAGAGGGGAGAG 421

RESULT 4
 AX556500 8354 bp DNA linear PAT 27-NOV-2002
 LOCUS Sequence 13 from Patent WO02057453.
 DEFINITION AX556500
 ACCESSION AX556500
 VERSION AX556500.1 GI:25899736
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyanankar,U.M.,
 Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E.,
 Zerhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.L.,
 Smithson,G., Li,L. and Ji,W.
 Polypeptides and nucleic acids encoding same
 Patent: WO 02057453-A 13 25-JUL-2002;
 Curagen Corporation (US)
 Location/Qualifiers
 1..8354
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

FEATURES
 source
 1..8354
 /location=Qualifiers
 1..8354

ORIGIN

Alignment Scores:
 Pred. No.: 7,51e-33 Length: 8354
 Score: 391.00 Matches: 76
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AX556500 (1-8354)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
 DB 7982 TCCAGATCAACACAGTACTTAATGCGAGGACTAGACGCTACACAGATCCAGCTCCAG 8041
 QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40

Mon Aug 16 09:01:13 2004

us-10-029-020-14_copy_2650_2725.p2n.rge

Db 8042 TACGGGCACTGTGCTTGAACACACGCTACCGGCAACCTTGATGAGCAGGACGCG 8101

Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60

Db 8102 GTCTTGGAGCTGGCCCGCAGAGAGCCCTGGCCCAAGCGTGGCCCGCAGCAGAGA 8161

Qy 61 LeuArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 76

Db 8162 CTGGGGGAAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8209

RESULT 5

AX675551

LOCUS

DEFINITION

AX675551

ACCESSION

AX675551.1

VERSION

AX675551.1

GI:29333552

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Padigaru, M., Li, L., Zerhusen, B.D., Casman, S.J., Shenoy, S., Spytek, K.A., Zhong, M., Gangolli, E.A., Burgess, C.E., Patturajan, M., Verne, C.A., Taylor, S., Tchernev, V.T., Miller, C.E., Guo, X., Boldog, F.L., Grosse, W.M., Alsobrook, J.P., Gerlach, V., Edinger, M., Rotherberg, M.E., Ellerman, K., Macdougall, J., Malyankar, U., Millet, I., Peyman, J., Smithson, G., Gunther, E. and Stone, D.J.

TITLE

Proteins, polynucleotides encoding them and methods of using the same

JOURNAL

Patent: WO 0205704-A 1 18-JUL-2002;

Curation Corporation (US)

FEATURES

source

1..8438

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 7.59e-33

Score: 391.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 6

US-10-029-020-14_COPY_2650_2725 (1-76) x AX675551 (1-8438)

Qy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20

Db 8033 TCCAGATCAACACAGTACTTATGCGAGGACTAGAGCTACACAGACATCCAGCTCCAG 8092

Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40

Db 8093 TACGGGCACTGTGCTTGAACACACGCTACGGGCAACGCTGGAGTGGAGGAGGAGGAGGAG 8152

Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60

Db 8153 GTCTTGGAGCTGGCCCGCAGAGAGCCCTGGCCCAAGCGTGGCCCGCAGCAGAGA 8212

Qy 61 LeuArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 76

Db 8213 CTGGGGGAAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8260

RESULT 6

AB037723

LOCUS

DEFINITION

AB037723

ACCESSION

AB037723.2

VERSION

AB037723.2

GI:20521827

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Nagase, T., Kikuno, R., Ishikawa, K.I., Hirose, M., and Ohara, O.

TITLE

Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro

JOURNAL

DNA Res. 7 (1), 65-73 (2000)

MEDLINE

20181126

PUBMED

10718198

REFERENCE

2 (bases 1 to 8624)

AUTHORS

Chara, O., Nagase, T. and Kikuno, R.

TITLE

Direct Submission

JOURNAL

Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdna@kazu.or.jp, URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913, Fax: +81-438-52-3914)

COMMENT

On May 9, 2002 this sequence version replaced gi:7242958.

FEATURES

Location/Qualifiers

1..8624

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_lib="pBCSKPCW"

/tissue="brain"

/note="This sequence was replaced that of fg06864 cDNA as a representative cDNA sequence for KIAA1302."

1..8624

/gene="KIAA1302"

/cdon_start=1

/product="KIAA1302 protein"

/protein_id="BAA92540.2"

/db_xref="GI:20521828"

/translation="STERLDNKPETTRVDSFGRLTNVPTQGVSSFRSDTSSVHVQ VETSKDDVTITNLSAGAFYLLQDQVRNSYVIGADGSLRLLLANGMEVALQTEPH LLAGTVNPTVGRNLTLPIDNGLNLEWQRKEQARQGVTVFGRRLRVHNRLLSLDF DRYTEKIDYDHRKEFRLILYDQAGRPSPSWRLNGVNVTVPGYIAGLQRLGM SERMEYDQAGRITRFADGKTWYFLEKSWLLSQRYIFEDNDKRLSSVTMP NVAROLETIRSVGYVNTVOPPEGNASVIQDFTEDGHLHLYFLGTGRRVYKYGKL SKLAETLYDTTKSVFTYDTAGMLKTNLNQNEGTCTIRYRQIGPLIDRQIFRTEEG MNARFDYNDNSFRVTSQAVINETPLDLYRVDVDSKTRQFGKFGVYIVANTDI ITTAVMTHTKFDAYGRMKVEQYEIFRSLMNTVQYDNNGRVVKELKVGYPANTTR YSYEYDADGQLTQVSIINDKELRYSYDLNGLNHLSPGNSARLPLRDLRDLRLGL DVQYKMEDEGFLRQGGDIPEYNSAGLLIKAYNRAGSWSVRYRVDGLGRVRSKSSHS HLQFFVADLTNPTKVTHLYNHSSEITSLYDLQHLFAMELSGDEFTIACDNIGT PLAVFSGTGIMIKILLYAYGEIVMTNPNFQIIIGYHGLYDPLTKLVHMRDIDYV LAGRWISPDHELKHLSSSNMPLNLYMKNPNPINSQDIKCFMTDVNSWLLTFGQ LHNVIQYKPKDMAMEPSTELIHTQKTOEWNSKSLIGVQCEVOKLAFVTLERF DQYLSGTTSCQAPKTKFASGSGVFGKGVKALKDGRVTTDIIISVANDGRVAIL LNHAYLNLHFTIDGVDTHYFVKPGSEGLALISGGRRTLENGVNVTSQINTV LNGRTRYTDIQLQYGCALCNTRYGTLDEKARVLELRQAVRAWEQRQLREG BEGLRAWTEGEKQQLVSTGRVQGVDFVSVQYPELSDSANNIHPMROSEMR"

ORIGIN

Alignment Scores:

Pred. No.: 7.77e-33

Score: 391.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 9

US-10-029-020-14_COPY_2650_2725 (1-76) x AB037723 (1-8624)

Qy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20

```

Db      3488  TCCAGATCAACAGACTTAATGTCAGCAGCTAGACGCTACACAGACATCCAGCTCCAG 3547
QY      21    TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGlyAlaArg 40
Db      3548  TAGGGGCACTGTGCTTGAACACACGCTACGGGCAACGCTTGATGAGGAGGACCG 3607
QY      41    ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db      3608  GTCTGTGAGCTGGCCCGGAGAGAGCCGTGCCAAGCCGTGGGCCCGGAGGAGGAGGAG 3667
QY      61    LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluGly 76
Db      3668  CTGCGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3715

RESULT 7
AX600210 LOCUS AX600210 8645 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Yue, H., Yao, M.G., Ison, C.H., Lu, Y., Warren, B.A., Elliott, V.S.,
Baughn, M.R., Ding, L., Xu, Y., Gietzen, K.J., Tang, T.Y., Lal, P.G.,
Duggan, B.N., Burford, N., Lu, D.A., Richardson, T.W., Tran, U.K.,
Khare, R. and Walla, N.K.
TITLE Proteins associated with cell growth, differentiation, and death
JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
INCYTE Genomics, Inc. (US)
FEATURES
source
1..8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7488573CH1"

ORIGIN
Alignment Scores:
Pred. No.: 7.79e-33 Length: 8645
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AX600210 (1-8645)
QY      1    SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db      8031  TCCAGATCAACAGACTTAATGTCAGCAGCTAGACGCTACACAGACATCCAGCTCCAG 8090
QY      21    TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGlyAlaArg 40
Db      8091  TAGGGGCACTGTGCTTGAACACACGCTACGGGCAACGCTTGATGAGGAGGAGGAGGAG 8150
QY      41    ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db      8151  GTCTGTGAGCTGGCCCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8210
QY      61    LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluGly 76
Db      8211  CTGCGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8258

RESULT 8
HSM806812 LOCUS HSM806812 8993 bp mRNA linear PRI 28-AUG-2003
DEFINITION Homo sapiens mRNA; cDNA DKF2p686K11107 (from clone DKF2p686K11107).
ACCESSION BX640737

```

```

VERSION BX640737.1 GI:34364828
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 8993)
Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
CONSRMT The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuerberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686K11107) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1..8993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="11q24.1"
/clone="DKFZp686K11107"
/tissue type="human fetal kidney"
/clone_lib="686 (synonym: hlcc3)". Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev stage="fetal"
1..8993
/gene="DKFZp686K11107"
<1..4200
/gene="DKFZp686K11107"
/note="hypothetical protein, N-terminus elongated,
differentially spliced"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAE45850.1"
/db_xref="GI:34364829"
/translation="RIDQNGIITLLGSLDLSARPLSCDSVMDISQVHLEWPTDLAI
NPWDNSLYLDNNVLIQISENHQVRIVAGRMHCQVPGIDHFLLSKVAIHAIHESALA
LAVSHNGVLIATDEKKNIRIQVTSGEISLVAGAPGCDCKNDANCSFGSDGQY
AKDAKLTSSSLAVCADGELYVADLGNIRIFIRKNKPFNTQNMELSSPIDQLYL
FDTTGKLYTQSLPTGTYLYNFTYTGDDITLTNNNGNMVNRDSTGMLVLPVD
GOVYVTMGTSNALKSVTQGHELAMTYHNSGLLATKSNENGWTTTFYEDSPGRLT
NVTFTQVSSFRSDTSSVHVQVETSSKDDVTITNLSAGAFYLLQDVRNSYII
GADGSLRLLANGMEVALQTEPHLLAGTVNPTVGNKRVLPIDNGLNLEWRQKQQA
RGQVTVGRRLRVHNRNLASLDPRVIRTEKIYDHRKFTLRILYDQAGPSLWSS
RLNGVNTSPGGIAGIQRMSEMEYDQAGRTSRIPADGKTWSYTYLLEKSMWII
LHSQRYIEFFDKNDRLSSVMPNVARQTLTIRSVGYRNIYQPEGNASVIOQDTE
DGHLLFTYLGTRRVIKYKGLSLAETLYDTTKVSTYVDETKMLKTINLQEGFT
CTIRYQIGLIDRLQIFRTEGMVNARFDYDNSFRVTSQCAVINETELPDLRY
DVGSGKTEQFGKGVLYIDNLIITAVMTHTFHFDAYGRMKEVQYIFRSLMWITV
QYDNMRVVKELKVGPIYANTRYSEYDADGQVTSINDKPLVSEYDNLNHLH
SPGNSRLVRIDRIDRITRLGDVQYKDEDFLFRQGGDIPEYNSAGLLIKAYNRA
GWSVRVYDGLGRRVSSSHHLOFFYADLTNPTKVTHLYNHSSETTSYLDLQ
GHLFAMELSSGDEFYIACVIGTAVFGTGLMIKQLLYTAYGETYMDPNDFOI11
GYHGGYLDPLTKLVHMRDNDYDLVAGRTSPDHELWKHLSSSNVPMFLMFKNNPI
SNSODIKCFMTDVNSWLLTFGQLHNVIPGYKPKDMEPESELLHTQMKTOEWDNS
KSLIGQCEVQQLKAFVTLERFDQYGSTITSQQAPKTKKFPASSGVSFGVGFAL
KGRVTDIISVANEDGRVVAALLNHAHLENLHFTIDGVDTHTFYKPGPSEGDALIL
GLSGRRITLNGVNTVQINTVNGTRRYTDIQLQYALCLNTRYGTTLDDEKARY
LELARQAVRQAWAREQRLREGEGLRATTEGEKQVQLSTGRVQGDGFFVYSVEQY
PELSDSANIHFMRSQEMGRR"
8953..8958
/gene="DKFZp686K11107"
8973
polyA_signal
polyA_site

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L12672
Center clone name: 886_F_14

ORIGIN
Alignment Scores:
Pred. No.: 8,148-33 Length: 8993
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x HSM806812 (1-8993)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 3838 TCCAGATCAACACAGTACTTAATGCAGGACTAGACGCTACACAGACATCCAGTCCAG 3897
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluysAlaArg 40
Db 3898 TAGCGGGCACTGTGCTTGAACACACGCTACGGGACACCTTGATGAGGAGGAGGCACGG 3957
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 3958 GTCCTGAGACTGCCCGGAGAGAGCGCGCCAGCGTGGCGCGGAGCAGCAGAGA 4017
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluys 76
Db 4018 CTGGGGAAGGGGAGGAGGAGCGCTGCGGGCTGACAGAGGGGGAGAAG 4065

RESULT 9
AC090378/c
LOCUS Homo sapiens chromosome 18 clone RP11-886F14 map 18, LOW-PASS
DEFINITION SEQUENCE SAMPLING.
AC090378
VERSION AC090378.1 GI:12958014
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 68970)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bouckhuyt, L., Bouckhuyt, B., Brown, A.,
Camarata, J., Campopiano, A., Choquel, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Gagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Headford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T.,
Lehotzky, J., Levine, R., Liu, G., MacLean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

* NOTE: This record contains 84 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 714: contig of 714 bp in length
* 814: gap of 100 bp
* 815 1552: contig of 738 bp in length
* 1553 1652: gap of 100 bp
* 1653 2377: contig of 725 bp in length
* 2378 2477: gap of 100 bp
* 2478 3210: contig of 733 bp in length
* 3211 3311: gap of 100 bp
* 3311 4036: contig of 726 bp in length
* 4037 4136: gap of 100 bp
* 4137 4837: contig of 701 bp in length
* 4838 4937: gap of 100 bp
* 4938 5678: contig of 741 bp in length
* 5679 6473: contig of 695 bp in length
* 6474 7277: contig of 704 bp in length
* 7278 8120: contig of 743 bp in length
* 8121 8221: gap of 100 bp
* 8221 8965: contig of 745 bp in length
* 8966 9066: gap of 100 bp
* 9067 9799: contig of 733 bp in length
* 9799 10591: contig of 693 bp in length
* 10592 10691: gap of 100 bp
* 10692 11413: contig of 722 bp in length
* 11414 12332: contig of 719 bp in length
* 12333 13067: contig of 735 bp in length
* 13068 13167: gap of 100 bp
* 13168 13898: contig of 731 bp in length
* 13899 14707: gap of 100 bp
* 14708 15514: contig of 707 bp in length
* 15515 16313: contig of 699 bp in length
* 16314 17161: gap of 100 bp
* 17162 17981: contig of 748 bp in length
* 17982 18081: contig of 720 bp in length
* 18082 18903: contig of 722 bp in length
* 18904 19633: gap of 100 bp
* 19634 20441: contig of 709 bp in length
* 20442 21272: contig of 731 bp in length
* 21273 21372: gap of 100 bp
* 21373 22089: contig of 717 bp in length
* 22090: contig of 717 bp in length

* 22090 22189: gap of 100 bp
* 22930 22934: contig of 745 bp in length
* 22935 23034: gap of 100 bp
* 23035 23777: contig of 743 bp in length
* 23778 23877: gap of 100 bp
* 24592 24592: contig of 715 bp in length
* 24693 25410: contig of 718 bp in length
* 25411 25510: gap of 100 bp
* 25511 26204: contig of 694 bp in length
* 26205 26304: gap of 100 bp
* 26305 27012: contig of 708 bp in length
* 27013 27112: gap of 100 bp
* 27113 27850: contig of 738 bp in length
* 27851 27950: gap of 100 bp
* 27951 28668: contig of 718 bp in length
* 28669 28768: gap of 100 bp
* 28769 29516: contig of 748 bp in length
* 29517 29616: gap of 100 bp
* 29617 30314: contig of 698 bp in length
* 30315 30414: gap of 100 bp
* 30415 31159: contig of 745 bp in length
* 31160 31259: gap of 100 bp
* 31260 31979: contig of 720 bp in length
* 31980 32079: gap of 100 bp
* 32080 32806: contig of 727 bp in length
* 32807 32906: gap of 100 bp
* 32907 33607: contig of 701 bp in length
* 33608 33707: gap of 100 bp
* 33708 34398: contig of 691 bp in length
* 34399 34498: gap of 100 bp
* 34499 35243: contig of 745 bp in length
* 35244 35343: gap of 100 bp
* 35344 36073: contig of 729 bp in length
* 36073 36172: gap of 100 bp
* 36173 36899: contig of 727 bp in length
* 36900 36999: gap of 100 bp
* 37000 37226: contig of 727 bp in length
* 37227 37826: gap of 100 bp
* 37827 38538: contig of 712 bp in length
* 38539 38638: gap of 100 bp
* 38639 39376: contig of 738 bp in length
* 39377 39476: gap of 100 bp
* 39477 40212: contig of 736 bp in length
* 40213 40312: gap of 100 bp
* 40313 41067: contig of 755 bp in length
* 41068 41167: gap of 100 bp
* 41168 41870: contig of 703 bp in length
* 41871 41970: gap of 100 bp
* 41971 42687: contig of 717 bp in length
* 42688 42787: gap of 100 bp
* 42788 43479: contig of 692 bp in length
* 43480 43579: gap of 100 bp
* 43580 44304: contig of 725 bp in length
* 44305 44404: gap of 100 bp
* 44405 45130: contig of 726 bp in length
* 45131 45230: gap of 100 bp
* 45231 45943: contig of 713 bp in length
* 45944 46043: gap of 100 bp
* 46044 46764: contig of 721 bp in length
* 46765 46864: gap of 100 bp
* 46865 47591: contig of 727 bp in length
* 47592 47691: gap of 100 bp
* 47692 48417: contig of 726 bp in length
* 48418 48517: gap of 100 bp
* 48518 49316: contig of 699 bp in length
* 49317 50042: contig of 726 bp in length
* 50043 50142: gap of 100 bp
* 50143 50882: contig of 740 bp in length
* 50883 50982: gap of 100 bp
* 50983 51700: contig of 718 bp in length
* 51701 51800: gap of 100 bp

* 51801 52548: contig of 748 bp in length
* 52549 52648: gap of 100 bp
* 52649 53362: contig of 714 bp in length
* 53363 53462: gap of 100 bp
* 53463 54172: contig of 710 bp in length
* 54173 54272: gap of 100 bp
* 54273 54997: contig of 725 bp in length
* 54998 55097: gap of 100 bp
* 55098 55817: contig of 720 bp in length
* 55818 55917: gap of 100 bp
* 55918 56444: contig of 727 bp in length
* 56445 56744: gap of 100 bp

Alignment Scores:

Pred. No.: 7,46e-32 Length: 68970
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AC090378 (1-68970)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 42327 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACAGACATCCAGCTCCAG 42268
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluysAlaArg 40
Db 42267 TACGGGCACTGTGCTTGAACACACGCTACGGGCAACGTTGGATGAGGAGGACGCG 42208
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrrAlaArgGluGlnGlnArg 60
Db 42207 GTCTCGAGCTGGCCCGCAGAGACGCTGCGCCAGCGTGGGCGGCGGCGGACGACGAGA 42148
QY 61 LeuArgGluGluGluGluGluGluLeuArgAlaTrrThrGluGluGluLys 76
Db 42147 CTGCGGGAAGGGGAGGAGGCGCTGCGGCGCTGGACAGAGGGGGGAGAG 42100

RESULT 10

AP002412/c

LOCUS

DEFINITION Homo sapiens chromosome 11 clone RP11-7H7 map 11g, WORKING DRAFT
SEQUENCE, 18 unordered pieces.

ACCESSION

AP002412

VERSION

AP002412.1 GI:8131676

KEYWORDS

HTG; HTGS PHASE1; HTGS_DRAFT.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 170156)

AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE

Published Only in DataBase (2000)

JOURNAL

2 (bases 1 to 170156)

AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE

Direct Submission

JOURNAL

Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Fax:81-42-778-9923, Fax:81-42-778-9924)

COMMENT

Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: <http://hgp.gsc.riken.go.jp/>
Contact: hattori@gsc.riken.go.jp
Project information
Center project name: HumDraft11

Center clone name: RP11-7H7

----- Summary Statistics -----

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 159074 bases at least Q40

Consensus quality: 164580 bases at least Q30

Consensus quality: 167092 bases at least Q20

Insert size: 168456; sum-of-contigs

Quality coverage: 5.57x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 18 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 29785 contig of 23056 bp in length

29886 52941 contig of 18687 bp in length

53042 71728 contig of 15844 bp in length

71829 87672 contig of 14301 bp in length

87773 102073 contig of 7249 bp in length

102174 109422 contig of 7900 bp in length

109523 117422 contig of 7389 bp in length

117523 124911 contig of 6147 bp in length

125012 131158 contig of 4489 bp in length

131259 135747 contig of 5700 bp in length

135848 141547 contig of 5406 bp in length

141648 147053 contig of 4376 bp in length

147154 151529 contig of 4743 bp in length

151630 156372 contig of 4695 bp in length

156473 161167 contig of 2835 bp in length

161268 164102 contig of 2835 bp in length

164203 167603 contig of 3401 bp in length

167704 170156 contig of 2453 bp in length.

* NOTE: This is a 'working draft' sequence. It currently consists of 18 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 29785: contig of 23056 bp in length

29886: gap of 100 bp

52941: contig of 23056 bp in length

53042: gap of 100 bp

71728: contig of 18687 bp in length

71829: gap of 100 bp

87672: contig of 15844 bp in length

87773: gap of 100 bp

102073: contig of 14301 bp in length

102174: gap of 100 bp

109422: contig of 7249 bp in length

109523: gap of 100 bp

117422: contig of 7900 bp in length

117523: gap of 100 bp

124911: contig of 7389 bp in length

125012: gap of 100 bp

131158: contig of 6147 bp in length

131259: gap of 100 bp

135747: contig of 4489 bp in length

135848: gap of 100 bp

141547: contig of 5700 bp in length

141648: gap of 100 bp

147053: contig of 5406 bp in length

147154: gap of 100 bp

151529: contig of 4376 bp in length

151630: gap of 100 bp

156372: contig of 4743 bp in length

156473: gap of 100 bp

161167: contig of 4695 bp in length

* 161168 161267: gap of 100 bp

* 161268 164102: contig of 2835 bp in length

* 164103 164203: gap of 100 bp

* 164203 167603: contig of 3401 bp in length

* 167604 167703: gap of 100 bp

* 167704 170156: contig of 2453 bp in length.

FEATURES

Location/Qualifiers

1..170156

source

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/map="llq"

/chromosome="11"

/clone="RP11-7H7"

1..29785

/note="assembly_fragment"

29886..52941

/note="assembly_fragment"

53042..71728

/note="assembly_fragment"

71829..87672

/note="assembly_fragment"

87773..102073

/note="assembly_fragment"

102174..109422

/note="assembly_fragment"

109523..117422

/note="assembly_fragment"

117523..124911

/note="assembly_fragment clone_end:T7 vector_side:left"

125012..131158

/note="assembly_fragment clone_end:SP6 vector_side:left"

131259..135747

/note="assembly_fragment"

135848..141547

/note="assembly_fragment"

141648..147053

/note="assembly_fragment"

147154..151529

/note="assembly_fragment"

151630..156372

/note="assembly_fragment"

156473..161167

/note="assembly_fragment"

161268..164102

/note="assembly_fragment"

164203..167603

/note="assembly_fragment"

167704..170156

/note="assembly_fragment"

ORIGIN

Alignment Scores:

Pred. No.: 1.99e-31 Length: 170156

Score: 391.00 Matches: 76

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 2 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AP002412 (1-170156)

Qy 1 SerGlnIleAenThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20

Db 116153 TCCAGATCAACACAGTACTTAAATGGCAGGACTAGCGCTACAGACATCCAGTCCAG 116094

Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40

Db 116093 TACGGGGCACCTGTCTTGAACACACGCTACGGGACACGTTGGATGAGAGAGCAGCG 116034

Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrrAlaArgGluGlnArg 60

Db 116033 GTCCTGGAGCTGGCGCGCAGAGAGCCGTGGCCCAAGCGTGGCGCCGAGCAGAGAGA 115974


```

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTThrGluGlyGluLys 76
|||||
Db 115973 CTGGCGGAGGAGGAGGAGGAGGCTGGCGGCTGGACAGAGGGGAGAG 115926

RESULT 11
AP002768/c
LOCUS
DEFINITION
Homo sapiens genomic DNA, chromosome 11q, clone:RP11-673F18,
complete sequence.
ACCESSION
AP002768
VERSION
AP002768.3 GI:13429926
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
PUBLISHED ONLY in Database (2000)
2 (bases 1 to 186084)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Mar 21, 2001 this sequence version replaced gi:11071944.
COMMENT
Location/Qualifiers
FEATURES
source
1. 186084
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-673F18"

ORIGIN
Alignment Scores:
Pred. No.: 2,2e-31 Length: 186084
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AP002768 (1-186084)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
|||||
Db 27964 TCCAGATCAACAGACTGACTTAATGGCAGACTAGACGCTACACAGACATCCAGCTCCAG 27905

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
|||||
Db 27904 TAGCGGCACTGCTTGTAACACACCGCTACCGGACACGCTGGATGAGGAGGACCGG 27845

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
|||||
Db 27844 GTCCGTGAGCTGCGCCGGCAGAGACGCTGCGCCAGCGTGGCCCGCAGAGCAGAGAG 27785

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
|||||
Db 27784 CTGGGAGAGGAGGAGGAGGAGGCTGGCGGCTGGACAGAGGGGAGAG 27737

RESULT 12
AK056531
LOCUS
DEFINITION
Homo sapiens cDNA FLJ31969 fis, clone NT2RP7008013, highly similar
to Mus musculus mRNA for Ten-m4.

```

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

AK056531
AK056531.1 GI:16551956
oligo capping; fis (full insert sequence).
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE AUTHORS

1 Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahari,K., Masuho,Y., Negai,K. and Isogai,T.
NEDO human cDNA sequencing project

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

2 (bases 1 to 2685)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; CDNA full insert sequencing:
Research Association for Biotechnology (RAB); CDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

FEATURES source

Location/Qualifiers
1. .2685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP7008013"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP7"
/notes="Cloning vector: pME18SPL3-mRNA from NT2 neuronal
precursor cells after 5-weeks retinoic acid (RA)
induction."
242. .2434
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71206.1"
/db_xref="GI:16551957"
/translations="MLKTNLQNEGFTTIRYQIGLIDRQIFRTEEGMVARFYY
NYDNSFRVTSQAVINETPLPIDLYEDVDSGTEQFGKGVYIDINQIITAVMTHT
TKHFDVAGMKVEQYEIFRSLMTVMVTQVDNMGKVKLKGVPYANTTRYSEYDAD
GQLQTSINDKPLWRYSDYDLNHLSPNSARLTPLRDYIDRDIRTLGDVQYKMD
DIEFGQRGDQIFEYNAGSLIKAYNRAGSVRYRYDGLRRVSSKSHHLOQFFVA
DLTNPKVTHLYNHSSEITSLYDIIQQLHFLAMELGSDGFYIACDNIGTAVFSGT
KIKQILYATAGEIYMDTNPNFQIIGVHGLYDPLTKLVHMRDQDYLAVGRWSP
DIELWKHLSSNMPENLYMFKNNPISNSQDIKCFMTDVNSWLLTFGQLHNVIPGY
PKPDMAMPSVELLHTQMTQEWKNSKILGVQCEVQKOLKAPFTLERFDQLYGSI
TSCLOPKTKTFASSGSVFGKFDKGRVTDDIISVANDGRVVAALNHAHYLE
NLHFTIDGVDDTHYFVKPGSEGLAILGLSGRRRTLENGVNTVTSQINTVLSQRTY
TDIQLQYALCLNTRYGTTLDEKARVLRLQRAVRQAWAREQQRLEGEGLRAWT
EGEKQQVLSTGRVQGVDFVTSVEQYPELSDSANNIHFWRQSEMGRR"

CDS

ORIGIN

Alignment Scores:
Pred. No.: 7,74e-33 Length: 2685
Score: 386.00 Matches: 75
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.68% Mismatches: 0
Query Match: 98.72% Indels: 0
DB: 9 Gaps: 0

Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 8.2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 6 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

* 1 103064: contig of 103064 bp in length
* 103065 103164: gap of unknown length
* 103165 124754: contig of 21590 bp in length
* 124755 124854: gap of unknown length
* 124855 147843: contig of 22989 bp in length
* 147844 147943: gap of unknown length
* 147944 168730: contig of 20787 bp in length
* 168731 168830: gap of unknown length
* 168831 192230: contig of 23400 bp in length
* 192231 192330: gap of unknown length
* 192331 208760: contig of 16430 bp in length.

FEATURES

source
1. .208760
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/clone="RP23-18J2"

ORIGIN

Alignment Scores:
Pred. No.: 8.38e-29 Length: 208760
Score: 368.00 Matches: 71
Percent Similarity: 97.37% Conservative: 3
Best Local Similarity: 93.42% Mismatches: 2
Query Match: 94.12% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AC074048 (1-208760)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 4946 TCCAGATCAACACCGTGCTCATGGCAGACATAGACGCTACACTGACATCCAGCTGCAG 5005
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 5006 TACAGGCGCTGTGCTGAACACCGCTACGGGACACACAGTGGATGAGGAAGTGGCG 5065
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 5066 GTGCTGAGCTGCCAGGAGAGAGAGTGTGCGCAGCTTGGCCGAGACGACGACAGA 5125
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 5126 CTGCGGGAAGGGAGAGAGCGCTCGGGCTGACAGATGGGAAAG 5173

RESULT 15

AC073599/c
LOCUS AC073599 226256 bp DNA linear ROD 28-FEB-2001
DEFINITION Mus musculus 7 BAC RP23-8M3 (Roswell Park Cancer Institute Mouse
BAC Library) complete sequence.
ACCESSION AC073599
VERSION AC073599.16 GI:13162472
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 226256)
AUTHORS Metzker,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C.,

Dederich,D., Thomas,S., Okwuonu,G., Carlock,C., Garner,T.,
Addison,S., Pace,A., Williams,G., Bonnin,D., Brooks,A., Brown,J.,
Buhay,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z.,
Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S.,
Fernandez,C., Ferraguto,D., Forcum-Tansey,J., Gill,R.,
Gorell,J.H., Gunaratne,P., Haller,G., Hernandez,J., Hogues,M.,
Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., Kelly,S.,
Kovar,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Martin,R.,
Massey,E., McLeod,M.P., Mei,G., Moore,S., Morgan,M., Morris,S.,
Neal,D., Nelson,A., Nguyen,R., Nguyen,N., Oguh,M., Parish,B.,
Perez,L., Reiter,D., Say,J., Shen,H., Vasquez,F., Watlington,S.,
Williamson,A., Wrenford,G., Zhou,X., Bouck,J., Hodgson,A.,
Munzy,M., Rives,M., Scherer,S., Sodergren,E., Weinstock,G.,
Worley,K. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 226256)
Worley,K.C.
Direct Submission
Submitted (26-JUN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 226256)
Worley,K.C.
Direct Submission
Submitted (28-FEB-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Feb 28, 2001 this sequence version replaced gi:12963002.
Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by
Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui
Zhang.

Exon/Intron boundaries of identified genes were chosen if there
were canonical splice junctions that maintained sequence continuity
across the splice junctions.

FEATURES

source
1. .226256
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/clone="RP23-8M3"
/note="Region similar to Mm#S1670873 601761001F1 Mus
musculus cDNA: BF124644"
938. .971
misc_feature /rpt_family="(CAAAA)n"
1118. .1165 repeat_region /rpt_family="T-rich"
complement(2209..2323) repeat_region /rpt_family="L2"
2468. .2493 repeat_region /rpt_family="(CAAA)n"
complement(2497..2613) repeat_region /rpt_family="B1-F"
3397. .3501 repeat_region /rpt_family="B4"
3758. .3948 repeat_region /rpt_family="B4"
5668. .5732 repeat_region /rpt_family="B4"
5733. .5809 repeat_region /rpt_family="PB1D10"
complement(5737..5848) repeat_region

```
repeat_region /rpt_family="YREP_MM"
repeat_region 5810..5820
repeat_region /rpt_family="B4"
repeat_region 6217..6254
repeat_region /rpt_family=" (TTGG)n"
repeat_region 6509..6566
repeat_region /rpt_family=" (TG)n"
repeat_region 6721..6756
repeat_region /rpt_family=" (CA)n"
repeat_region complement(7827..7942)
repeat_region /rpt_family="PB1D10"
repeat_region complement(7899..7964)
repeat_region /rpt_family="B4"
repeat_region complement(7965..8038)
repeat_region /rpt_family="ID4"
repeat_region 8093..8191
repeat_region /rpt_family="L2"
repeat_region 8326..8437
repeat_region /rpt_family="MIR"
repeat_region 9579..9616
repeat_region /rpt_family=" (CATA)n"
repeat_region 9641..9664
repeat_region /rpt_family=" (CCA)n"
repeat_region complement(9715..9783)
repeat_region /rpt_family="PB1D10"
repeat_region complement(9778..9830)
repeat_region /rpt_family="ID3"
repeat_region 9927..9958
repeat_region /rpt_family=" (CAAA)n"
repeat_region complement(10121..10237)
repeat_region /rpt_family="B1-F"
repeat_region 10243..10414
repeat_region /rpt_family=" (CCTG)n"
repeat_region 10440..10509
repeat_region /rpt_family="CT-rich"
repeat_region complement(10532..10578)
repeat_region /rpt_family="B1-F"
repeat_region 10971..11355
repeat_region /rpt_family="Lx6"
misc_feature complement(11571..13187)
misc_feature /notes="Region similar to Mm#S685488 Odd Oz/ten-m homolog 4 (Drosophila): AB025413"
repeat_region 14505..14614
repeat_region /rpt_family="PB1"
repeat_region 15111..15194
repeat_region /rpt_family="MIR"
repeat_region complement(15863..16043)
repeat_region /rpt_family="B4"
repeat_region 16385..16459
repeat_region /rpt_family="MIR"
repeat_region 19475..19693
repeat_region /rpt_family="B2"
repeat_region 20006..20169
repeat_region /rpt_family="B3"
repeat_region 20170..20190
repeat_region /rpt_family=" (TA)n"
repeat_region 20191..20251
repeat_region /rpt_family="B3"
repeat_region complement(21402..21484)
repeat_region /rpt_family="MIR"
repeat_region 24484..24563
repeat_region /rpt_family="L2"
repeat_region 24789..24907
repeat_region /rpt_family="L2"
repeat_region complement(25094..25241)
repeat_region /rpt_family="MIR"
repeat_region complement(26937..27347)
repeat_region /rpt_family="L1MCC"
repeat_region 27569..27613
repeat_region /rpt_family=" (TTTA)n"
repeat_region complement(28594..28633)
repeat_region /rpt_family="L2"
repeat_region complement(28841..28955)
```

```
repeat_region /rpt_family="PB1D10"
repeat_region 29428..29480
repeat_region /rpt_family=" (TATG)n"
repeat_region 29510..29582
repeat_region /rpt_family=" (CAGA)n"
repeat_region 29583..29723
repeat_region /rpt_family="GA-rich"
repeat_region 29776..29959
repeat_region /rpt_family=" (CAGAGA)n"
repeat_region complement(30193..30338)
repeat_region /rpt_family="B1_MM"
repeat_region 31182..31452
repeat_region /rpt_family="B4"
repeat_region 31913..31950
repeat_region /rpt_family="AT-rich"
repeat_region 32864..32885
repeat_region /rpt_family="AT-rich"
repeat_region complement(33023..33070)
repeat_region /rpt_family="MIR"
repeat_region 34393..34458
repeat_region /rpt_family="MERSB"
repeat_region complement(34687..34859)
repeat_region /rpt_family="MERSA"
repeat_region 34872..34950
repeat_region /rpt_family="MERSA"
repeat_region 35660..35869
repeat_region /rpt_family="URR1A"
repeat_region 36089..36106
repeat_region /rpt_family="B4"
repeat_region 36107..36239
repeat_region /rpt_family="B1_MM"
repeat_region 36263..36295
repeat_region /rpt_family=" (CCA)n"
repeat_region 36398..36419
repeat_region /rpt_family=" (CCCCAG)n"
repeat_region 36790..36817
repeat_region /rpt_family=" (T)n"
repeat_region complement(38755..38926)
repeat_region /rpt_family="B4"
repeat_region complement(39294..39403)
repeat_region /rpt_family="L2"
repeat_region 40007..40038
```

Alignment Scores:
Pred. No.: 9.15e-29 Length: 226256
Score: 368.00 Matches: 71
Percent Similarity: 97.37% Conservative: 3
Best Local Similarity: 93.42% Mismatches: 2
Query Match: 94.12% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AC073599 (1-226256)

```
Qy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 1609 TCCAGATCAACACCGTGTCTAGTGCAGGACTAGACGCTACACTGCAGCTGCAG 1550
Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluLysAlaArg 40
Db 1549 TACAGGGCGTGTGCTGAACACCCGCTACGGGACACAGTGGATGAGGAAAGTGC 1490
Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgAlaAlaTrpAlaArgGluGlnArg 60
Db 1489 GTGCTGGAGCTGGCCAGGACAGAGCTGTGCGGCGCTTGGGCCCGGAGGACGAGA 1430
Qy 61 LeuArgGluGlyGluGluGlyLeuArgAlaAlaTrpThrGluGlyGluLys 76
Db 1429 CTGGGGAAGGGGAAGAGGGCGCTTGGCGGCTTGCACAGATGGGGAAG 1382
```

Search completed: August 14, 2004, 12:05:34
Job time : 1496.2 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 14, 2004, 19:26:18 ; Search time 602 Seconds

(without alignments)

3260.204 Million cell updates/sec

Title: SEQ14-X-AT-28-64-76

Perfect score: 2132

Sequence: 1 MDVKERKPYRSLRRDAER.....EITDTASSWPVPTDVSLLP 400

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/cgn2_1/USFO.spool/MITRA020/runat_06082004.114103.258/app_query.fasta_1.583
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom2
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=MITRA020 @CIGN 1.1.519 @runat 06082004.114103.258
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/2/pubpna/US0E_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/2/pubpna/PTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
|------------|-------|-------------|--------|----|-------------|

ALIGNMENTS

RESULT 1
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

| | | | | | | |
|----|-------|------|-------|----|--------------------|--------------------|
| 1 | 2127 | 99.8 | 8354 | 13 | US-10-383-201-43 | Sequence 43, Appl |
| 2 | 2127 | 99.8 | 8354 | 13 | US-10-029-020-13 | Sequence 13, Appl |
| 3 | 2122 | 99.5 | 1680 | 15 | US-10-290-578-3 | Sequence 3, Appl |
| 4 | 2122 | 99.5 | 3111 | 9 | US-09-773-517-12 | Sequence 12, Appl |
| 5 | 2122 | 99.5 | 3111 | 9 | US-09-792-025-12 | Sequence 12, Appl |
| 6 | 2122 | 99.5 | 3111 | 9 | US-09-849-868-12 | Sequence 12, Appl |
| 7 | 2122 | 99.5 | 3111 | 15 | US-10-290-578-1 | Sequence 1, Appl |
| 8 | 2122 | 99.5 | 3111 | 15 | US-10-453-183-12 | Sequence 12, Appl |
| 9 | 2114 | 99.2 | 8355 | 13 | US-10-383-201-55 | Sequence 55, Appl |
| 10 | 2082 | 97.7 | 8438 | 13 | US-10-042-865-1 | Sequence 1, Appl |
| 11 | 1285 | 60.3 | 2387 | 15 | US-10-290-578-11 | Sequence 11, Appl |
| 12 | 975 | 45.7 | 8645 | 17 | US-10-038-854-37 | Sequence 37, Appl |
| 13 | 975 | 45.7 | 8675 | 17 | US-10-038-854-35 | Sequence 35, Appl |
| 14 | 937 | 43.9 | 9695 | 16 | US-10-144-194A-81 | Sequence 81, Appl |
| 15 | 905 | 42.4 | 2496 | 9 | US-09-808-602-75 | Sequence 75, Appl |
| 16 | 905 | 42.4 | 2496 | 10 | US-09-800-198-63 | Sequence 63, Appl |
| 17 | 904 | 42.4 | 8689 | 9 | US-09-808-602-78 | Sequence 78, Appl |
| 18 | 904 | 42.4 | 8689 | 10 | US-09-800-198-66 | Sequence 66, Appl |
| 19 | 904 | 42.4 | 8737 | 9 | US-09-808-602-74 | Sequence 74, Appl |
| 20 | 904 | 42.4 | 8797 | 9 | US-09-808-602-77 | Sequence 77, Appl |
| 21 | 904 | 42.4 | 8797 | 10 | US-09-800-198-62 | Sequence 62, Appl |
| 22 | 904 | 42.4 | 8797 | 10 | US-09-800-198-65 | Sequence 65, Appl |
| 23 | 885 | 41.5 | 8409 | 9 | US-09-808-602-79 | Sequence 79, Appl |
| 24 | 885 | 41.5 | 8409 | 10 | US-09-800-198-67 | Sequence 67, Appl |
| 25 | 859 | 40.3 | 8575 | 13 | US-10-072-012-143 | Sequence 143, Appl |
| 26 | 828 | 38.8 | 9729 | 9 | US-09-808-602-12 | Sequence 12, Appl |
| 27 | 828 | 38.8 | 9729 | 10 | US-09-800-198-12 | Sequence 12, Appl |
| 28 | 828 | 38.8 | 9826 | 9 | US-09-808-602-7 | Sequence 7, Appl |
| 29 | 828 | 38.8 | 9826 | 10 | US-09-800-198-7 | Sequence 7, Appl |
| 30 | 733.5 | 34.4 | 12880 | 16 | US-10-295-027-927 | Sequence 927, App |
| 31 | 641.5 | 30.1 | 8473 | 17 | US-10-038-854-39 | Sequence 39, Appl |
| 32 | 641.5 | 30.1 | 8487 | 17 | US-10-038-854-41 | Sequence 41, Appl |
| 33 | 522 | 24.5 | 1755 | 10 | US-09-398-966-17 | Sequence 17, Appl |
| 34 | 522 | 24.5 | 1755 | 15 | US-10-004-415-17 | Sequence 17, Appl |
| 35 | 522 | 24.5 | 1755 | 16 | US-10-384-974-17 | Sequence 17, Appl |
| 36 | 508 | 23.8 | 1727 | 16 | US-10-094-749-710 | Sequence 710, App |
| 37 | 508 | 23.8 | 9058 | 16 | US-10-144-194A-79 | Sequence 79, Appl |
| 38 | 452 | 21.2 | 527 | 15 | US-10-029-386-1356 | Sequence 1356, Ap |
| 39 | 434 | 20.4 | 1430 | 9 | US-09-808-602-5 | Sequence 5, Appli |
| 40 | 434 | 20.4 | 1430 | 10 | US-09-800-198-5 | Sequence 5, Appli |
| 41 | 434 | 20.4 | 1431 | 10 | US-09-977-418-21 | Sequence 21, Appl |
| 42 | 434 | 20.4 | 1431 | 10 | US-09-977-033A-21 | Sequence 21, Appl |
| 43 | 434 | 20.4 | 1431 | 10 | US-09-977-751C-21 | Sequence 21, Appl |
| 44 | 434 | 20.4 | 1431 | 10 | US-09-977-639A-21 | Sequence 21, Appl |
| 45 | 434 | 20.4 | 1431 | 11 | US-09-977-819B-21 | Sequence 21, Appl |

8354

Db 1 ATGGACGTGAAGAGAGAGAGCCCTTACCGCTCGCTAGCCGCGCGCGAGCGCCGAGCGC 60
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 61 CGCTACACCACTCGTCCCGGACAGCAGCGAGGAGGCAAGCGCCGACGAAATCGTACAGC 120
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 121 TCCAGCAGACCCCTGAAGCCCTACACACGAGACGCCCGCTAGCCTATGCGACCGCGCTC 180
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 181 AAGGACATGTGCGCAGAGGCCCGAGGAAATCTCGCGACAGGTGCCAATCTCACCTG 240
QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 241 CGGAGCTGGGGCTGGAAGGTAACCGCCCTCACCGGACCCCTGTACCGGACAGACATT 300
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 301 GGCCCTCCCGCCACTGGGGCTACTCCATGGGGGCTGGCTCTGATGCCACATGGAGGCTGAC 360
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 361 ACGGTGCTGTCCCTGAGCACCCCGTGGCTCTGTGGGCGCGAGCACACGGTCAGGGCC 420
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 421 AGCTCTGCTGCTCCAGCGGGCCAAATTCCTCACTCACCTCACCGACACCGAGCATGAA 480
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 481 AACACTGAGACTGATCATCCGGCGGCGCTGCAGAACCCAGCGGGCTCCGGAGCGCGCCG 540
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaSerIleAsnSerLeu 200
Db 541 CGCGCGCTCTCGACGCCCCACACCCCAACAGCAGCACCGCGGCTCCATTAACCTCCCTG 600
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 601 AACCGGGGCAACTTACGCGGAGGAGCAACCCCGAGCCCGCGCCCGCCACCGACCACTCGCTC 660
QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 661 TCCGAGAGCCCTCTCCCGCGCGCGCCAGAGGCTTGCACCCCGCCAGAGAACTGGCTG 720
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 721 CTCAACAGCAACATCCCTCGGAGACAGAAACCTAGGCAAGCAGCAATTCCTTAGGGACA 780
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 781 TTGACAGCAACCTCATGAGATGGACATTCCTCGCGCTCTCCGCTATGATGGGCTTAC 840
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 841 AGTGAGGGGCACTTCCTCTCAAGCTGGAGGACCTCCCGCTCTCTTCTGCACCATCA 900
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 901 CCAGGGTACCACTGACGCTCCAGCACAGTGTACTCTCTCCGCGCCGAGCCCTTCGCCCGC 960
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 961 AGCACTCTGCCCGCGCGCTTAACTCAAGAGCCCTCAAGTACTGTAAGTACTGGAAG 1020
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1021 TGGCGAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGTGTCATCTCTGCTGGCATACTTT 1080
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

Db 1081 GTGGCCATGCACCTGTTTGGCTAACTGGCACCTGCAGCCGATGAGGGGCGAGATGTAT 1140
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1141 GAGATCAGGAGGACACAGCCAGCAGTGGCTGTGGCCACCGACGCTCTCCCTATACCCC 1200
RESULT 4
US-09-773-517-12
; Sequence 12, Application US/09773517
; Patent No. US20010023241A1
; GENERAL INFORMATION:
; APPLICANT: Sliwkowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; FILE REFERENCE: P1145R1
; CURRENT APPLICATION NUMBER: US/09/773,517
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/243,198
; PRIOR FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-773-517-12
Alignment Scores:
Pred. No.: 5, 07e-206 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 9 Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x US-09-773-517-12 (1-3111)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGACGTGAAGAGAGAGGAAAGCTTACCGCTCGCTAGCCGCGCGCGAGCGC 393
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCACTCGTCCCGGACAGCAGGAGGCGCAAGCGCCGAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCAGACCCCTGAAGCCCTACGACAGAGACCGCCGCTAGCCTATGCGACCGCGCTC 513
QY 61 LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db 514 AAGGACATTTGTCCGCGAGGCGCGAGGAAATTCGCGCGACAGGTGCCAATTCACCCCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGGCTGGAAGAGTAACGCCCTCACCGGACCCCTGTACCGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCCTCCCGCCACTGGCGCTACTCCATGGGGGCTGGCTCTGATGCCACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTGAGCACCCCGCTGTGTGGGCGCGAGCACCGGTGAGGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTGCTGCTCCAGCGGGCCAAATTCCTCACTCACCTCACCGACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGCGGCGCTGCAGAACCGACGCGCGGTCCGAGCGCGCCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaSerIleAsnSerLeu 200

| | | | |
|----|------|---|------|
| Db | 874 | CGCGCGTCTCGACGGCCACACCCCAACAGACACACGGCGGCTCCATTAACTCCCTG | 933 |
| Qy | 201 | AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu | 220 |
| Db | 934 | AAACGGGGCAACTTCAGCCGAGGAGCAACCCAGCCCGGGCCCAACGGACCATCGGTC | 993 |
| Qy | 221 | SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu | 240 |
| Db | 994 | TCCGGAGAGCCCTCGCGGGCGGCCAGAGAGCTGCCACGCCACGAGAACTGGGTG | 1053 |
| Qy | 241 | LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr | 260 |
| Db | 1054 | CTCAACAGCAACATCCCTCGAGACCAGAAACCTAGCAAGCAGCCATCTCTAGGGACA | 1113 |
| Qy | 261 | LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr | 280 |
| Db | 1114 | TTGCAGAGACAACCTCATGTAGATGGACATCTCGCGCCCTCCGCCCATGATGGGGCTTAC | 1173 |
| Qy | 281 | SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer | 300 |
| Db | 1174 | AGTGACGGGCATTCCTCTTCAACGCTGGAGGCACCTCCCGCTCTCTCGACCAATCA | 1233 |
| Qy | 301 | ProGlyTyrProLeuThrSerSerThrValTyrSerProProProProLeuProArg | 320 |
| Db | 1234 | CCAGGGTACCCACTGACGTCCAGCACAGTGACTCTCTCGCGCCGACCCCTGCCCGCG | 1293 |
| Qy | 321 | SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys | 340 |
| Db | 1294 | AGCACCTTCGCCGGCGGGCTTTAACTCAAGNAGCCCTCCAAGTACTGTAACTGGAG | 1353 |
| Qy | 341 | CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe | 360 |
| Db | 1354 | TGCGACGCCCTGAGCGGCATCGTCATCTCAGCCACTCTGTGTCACTCTCGTGGCATACTTT | 1413 |
| Qy | 361 | ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr | 380 |
| Db | 1414 | GTGGCCATGCACCTGTTTGCGCTAACTGGCACTTCGACCGCATGGAGGGCGAGATGTAT | 1473 |
| Qy | 381 | GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro | 400 |
| Db | 1474 | GAGATCACGGAGACACAGCACAGTGGCCCTGTGCACACCGACGTCCTCCCTATACCCC | 1533 |

RESULT 5

US-09-792-025-12
Sequence 12, Application US/09792025
Patent No. US20020042087A1
GENERAL INFORMATION:
APPLICANT: Sliwowski, Mark
Kern, Jeffrey
TITLE OF INVENTION: Use of Heregulin as a Growth Factor
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/792,025
FILING DATE: 23-Feb-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/020,598
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Schwartz, Timothy R.

```
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACACAGACATCCCTCGAGACCAAGAACCTTAGCAGACGCCATTTCTTAGGAGACA 1113
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGACGAGACACCTCATTCAGATGACATTCCTGGCGCTCCCGCCATGATGGGCTTAC 1173
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCAATCA 1233
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProProProProArg 320
Db 1234 CAGGGTACCCACTGACGTCCAGCAGATGACTCTCTCCGCCCGCACCCCTGCCCCGC 1293
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCCCGCGCCCTTTAACTCAAGAGCCCTCCAGTACTGTAACTGGAAG 1353
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCAGGCGCTGAGCGCCATCTCATCTCAGCCACTCTGGTCTATCTCTGTCGCACTTT 1413
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTTGGCTTAACTGGACCTGACCCGATGGAGGGGCGAGATAT 1473
Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACACAGCCAGCAGTTGGCTGTGCCAACCGACGTCTCCCTATACCC 1533
```

RESULT 6

```
US-09-849-868-12
; Sequence 12, Application US/09849868
; Patent No. US20020081299A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Gao, Wei-Qiang
; TITLE OF INVENTION: HAIR CELL DISORDERS
; FILE REFERENCE: GENENT.035VPC
; CURRENT APPLICATION NUMBER: US/09/849,868
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (334)...(2637)
US-09-849-868-12
```

```
Alignment Scores:
Pred. No.: 5,07e-206 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: Gaps: 0
```

SEQ14-X-AT-28-64-76 (1-400) x US-09-849-868-12 (1-3111)

```
Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGACGCTGAAGAGAGAGAGCCCTTACCGCTCGCTGACCCGCGCGCGAGCGGAGCGC 393
Qy 21 ArgTyrThrSerSerAla***SerGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCAAGCTCGTCCGCGGACGAGAGGGGCAAGCCCGCGAGAAATCGTACAGC 453
Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
```

```
Db 454 TCCAGCAGAGACCCCTGAAGGCCCTAGCAGCCAGGACGCCCGCTAGCCTATGGCAGCGCGCTC 513
Qy 61 LysAspIle***ProGlnGluAlaGluPheCysArgThrGly**AsnPheThrLeu 80
Db 514 AAGACACATTTGCGCGCAGGAGCGCGAGAAATTCGCCGACAGGTGCCACTTACCCCTG 573
Qy 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGGCTGGAAGAGTAACGCCCTCACGGGACCCCTGTACCGGACAGACATT 633
Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCGCCACTGCGGCTACTTCCATGGGGGCTGCTCTGATGCCGACATGGAGGCTGAC 693
Qy 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTTGAGCACCCCGTGGCTCTGTGGGGCGGAGACACGCGTCAGGGCGC 753
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTCTGCTTCCAGCGCGGCCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGCGCGCTGCAGAACACACGCGCGCTCCGGACCGCGCG 873
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
Db 874 CGCGCGCTCTCCAGCGCCACACCCCAACACAGCACCGCGCGCTCCATTAACCTCCCTG 933
Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGGCACTTTCAGCGCGAGGAGCAACCCAGCGCGCGCCCGACCGACCACTCGCTC 993
Qy 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGAGAGAGCCCGCTCCCGCGCGCGCCAGGAGCCTTCCCGCGCGCTCCCGCGCGCTTAC 1053
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACACAGACATCCCTCGAGACCAAGAACCTTAGCAGACGCCATTTCTTAGGAGACA 1113
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGACGAGACACCTCATTCAGATGACATTCCTCGCGCGCTCCCGCCATGATGGGCTTAC 1173
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCAATCA 1233
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProProProProArg 320
Db 1234 CAGGGTACCCACTGACGTCCAGCAGATGACTCTCTCCCGCGCGCGCGCGCGCGCTTAC 1293
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCCCGCGCCCTTTAACTCAAGAGCCCTCCAGTACTGTAACTGGAAG 1353
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCAGGCGCTGAGCGCCATCTCATCTCAGCCACTCTGGTCTATCTCTGTCGCACTTT 1413
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTTGGCTTAACTGGACCTGACCCGATGGAGGGGCGAGATGAT 1473
Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACACAGCCAGCAGTTGGCTGTGCCAACCGACGTCTCCCTATACCC 1533
```

RESULT 7

US-10-290-578-1

```

; Sequence 1, Application US/10290578
; Publication No. US20030078389A1
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; Sliwkowski, Mark
; TITLE OF INVENTION: Gamma-Heragulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/290,578
; FILING DATE: 08-No. US20030078389A1-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/021640
; FILING DATE: 12-Jul-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-290-578-1

Alignment Scores:
Pred. No.: 5.07e-206 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 15 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-10-290-578-1 (1-3111)

QY 1 MetAspValLysGluArGlySerProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
Db 334 ATGGACGTTGAAGAGAGAGAGAGAGCCATTACCGCTCGCTGACCCGCGCGCGAGCGCG 393
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCGCTCTGTCGCGGAGAGGAGGAGGCGCAAGCCCGCGCAAAATCGTACAG 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCTTGAAGCCCTACGACCAGACGCCCGCTTACGCTATGCGACGCCGCTG 513
QY 61 LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db 514 AAGGACATTGTGCGCAGGAGCGGAGAAATTCGCGGCACAGGTGCGCAACTTACCCCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100

```

;; PRIOR APPLICATION NUMBER: US/09/243,198
;; PRIOR FILING DATE: 1999-02-02
;; PRIOR APPLICATION NUMBER: US 60/073,866
;; PRIOR FILING DATE: 1998-02-04
;; NUMBER OF SEQ ID NOS: 14
;; SEQ ID NO 12
;; LENGTH: 3111
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-453-183-12

Alignment Scores:
Pred. No.: 5,07e-206 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservatives: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 15 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-10-453-183-12 (1-3111)

QY 1 MetAspValysGluArgGlyProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGACGTGAAGAGAGAAAGCTTACCGCTCGCTGACCCGCGCGCGAGCCGAGCGC 393

QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACACGCTCGTCGCGGACGAGCGAGGAGGCAAGCCCGCGAATCTGTACAGC 453

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCACGGAGACCTGAGCGCTACGACGAGCGCGCGCTAGCTAGCTATGGACGCGCTC 513

QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 514 AAGGACATTTGTCGCGAGGCGCGAGGAATTTCTGCGCACAGGTGCCAACTTCACCTG 573

QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGAGACCTGTACCGGACAGCAT 633

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCCACTGCGGCTACTCCATGGGGGCTGCTCTGATGCGGACATGAGGCTGAC 693

QY 121 ThrValLeuSerProGluHisProValArgLeuTyrPheGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTGAGCACCCCGTGTCTGTGGGGCGGAGCACACGGTCAGGGCGC 753

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTGCTGTCCAGCGGGGCAATTCATCTCACACTCCCGACACCGAGCATGAA 813

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGCGCGCTGCAGAACACCGCGGCTCCGGAGCGCGCG 873

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CGCGCGCTCTCGCACGCCACACCCCAACCCAGCACCGCGGCTCCATTAATCTCCCTG 933

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGGCACTTTCACGCGGAGAGCAACCCCGCGCGCGCCCGCGGAGCTCGCTC 993

QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnThrLeu 240
Db 994 TCCGGAGAGCCCTGCGCGCGCGCGCCAGAGGCTGCCACACCGCCCGAGGAGAACTGCTG 1053

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCAACAGCAACATCCCTCTGGAGACAGAAACCTTAGGCAAGCAACCATCTCTAGGACA 1113

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACAACTTCATTGAGATGACATTCCTCGCGCGCTCCCGCCCATGATGGGGCTTAC 1173

QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTCAGCGGCACTTCCTCTTCAAGCCTGGAGGACCTCCCGCGCTCTCTGACACCATCA 1233

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGATACCCACTGACGCTCCAGCACAGTACTCTCTCTCCGCGCGACCCCTGCGCGC 1293

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTyrLys 340
Db 1294 AGCACCCTTCGCGCGCGCGCTTTAACTCAAGAGCGCTCCCAAGTACTGTAACCTGGAAG 1353

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TCGCAGCGCCCTGAGCGCATCGTCATCTCAGCCACTCTGGTCATCTCTGCTGGCATCTT 1413

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTTTGGCTTAACTGGCACCTGCAGCGCATGAGAGGGGAGATGAT 1473

QY 381 GluLeuThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCAGTTGGCTGTGTGCAACCGACGTCTCCCTATACCC 1533

RESULT 9
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:
Pred. No.: 1.24e-204 Length: 8355
Score: 2114.00 Matches: 395
Percent Similarity: 98.75% Conservatives: 0

Percent Similarity: 98.01% Conservative: 0
Best Local Similarity: 98.01% Mismatches: 5
Query Match: 97.65% Indels: 3
DB: 13 Gaps: 1

SEQ14-X-AT-28-64-76 (1-400) x US-10-042-865-1 (1-8438)

```
Qy 1 MetAspValIysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 4 ATGGACGTGAAGGAGAGAGACCTTACCGCTCGCTGACCCGCGCGCGGAGCGCGAGCGC 63
Qy 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 64 CGCTACACCACTGCTCGCGGAGCAGCGAGGAGGCAAGCCCGCGCAATCGTACAGC 123
Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 124 TCCAGCGAGACCTTGAAGGCTTACGACCGAGCGCGCTAGCCTATGCGAGCGCGCTC 183
Qy 61 LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db 184 AAGGACATTTGTCGCGCAGGAGCGCGAGAAATCTGCCGACAGGTCGCACTTCCACCTG 243
Qy 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 244 CGGGAGCTGGGGTGAAGAGTAGTACGCCCTCTACGGGACCTGTACGGGACAGACAT 303
Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 304 GCGCTCCCCCACTGCGGGTACTCCATGGGGGTGGCTCTGTATGCCGACATGGAGGCTGAC 363
Qy 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 364 ACGGTGCTGTCCCTTGTAGCACCCCGGTGCTGTGGGCGGAGGACACCGTCAAGGCGC 423
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 424 AGCTCTCGCTGCTCCAGCGGGGCAATCCATCTCACACTCACCGACACCGGACATGAA 483
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 484 AACACTGAGACTGATCATCCGGCGGCGCTGCAGAACACCGCGGCTCCGGAGCGCGCG 543
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
Db 544 CCGCGGCTCTCGACGCCACACCCCAACAGCACCGCGGCTCCATTAACCTCCCTG 603
Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 604 AACCGGGCACTTCCAGCGGAGGAGCAACCCCGCGGCGCGCCACCGGACCACTCGCTC 663
Qy 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTyrLeu 240
Db 664 TCCGAGAGCCCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 723
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 724 CTCACAGCAACATCCCTCTGAGACACAGAACCTAGGACAGAGCACTCTCTAGGACA 783
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 784 TTGACGACCAACCTCATTCAGATGGACATTCGCGGCGCTCCCGCATGATGGGCGCTAC 843
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 844 AGTGACGGGACCTTCTCTTCAAGCTGGAGGACCTCCCGCTCTCTTCGCACCATCA 903
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 904 CCAAGGTACCACTGACGTCCAGCACAGTACTCTCTCCGCGCGCGCGCGCGCGCGCG 963
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTyrLys 340
```

```
Db 964 AGCACCTTCGCTGCGCGGCTTTAACTCAAGAGCCCTCCAAGTACTGTAACTGGAAG 1023
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1024 TCGCAGCCCTGAGCGGCATCGTCACTCAGCCACTCTGGTCACTCTGCTGGCATCTTT 1083
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet--- 379
Db 1084 GTGGCCATGCACCTGTTTGGCTAAACTGGCACCTGCAGCCGATGGAGGGCGCATGACG 1143
Qy 380 ----TyrGluLeuThrGluAspThrAlaSerSerTrpProValProThrAspValSerLe 398
Db 1144 GATTATGAGATCACGGAGGACACAGCCAGCAGTTGGCTGTGCAACCGACGCTCCCT 1203
Qy 398 uTyrPro 400
Db 1204 ATACCC 1210

RESULT 11
US-10-290-578-11
; Sequence 11, Application US/10290578
; Publication No. US20030078389A1
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herectulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/290,578
; FILING DATE: 08-Nov-2003/078389A1-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/021640
; FILING DATE: 12-Jul-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-290-578-11
```

Alignment Scores: 8,31e-121 Length: 2387
Pred. No.: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.27% Indels: 0
DB: 15 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-10-290-578-11 (1-2387)

```
QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGCGCTCGAGAACACCGCGCGCTCGGAGCGCGCGCGCTCGCAC 60
QY 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCACACCCCCAACAGACAGCACCGCGCGCTCCATTAACTCCCTGAACCGGGGCACTTC 120
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACGCCGAGGAGCAACCCAGCGCGCGCCCGACGACCTCGCTCTCGGAGAGCCCCCT 180
QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
Db 181 GCGCGCGCGCGCCAGGAGCGCTGCCACCGCCAGGAGAACTGGCTGTCTCAACAGCAACATC 240
QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
Db 241 CCCCTGGAGACCAAGAACTAGCAGACGACCATTCCTAGGACATTCGAGGACAACTC 300
QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGACATTTCTCGCGCGCTCCCGGCATGATGGGCTTTACAGTGACGGGCACTTC 360
QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 361 CTCTTCAAGCCTGGAGGACCTCCCGCGCTCTTCGACCAACATCACCAGGGTACCCTG 420
QY 306 ThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 AGCTCCAGCACAGTGACTCTCTCCGCCCGGACCTCGCCCGAGACCTTCGCGCCGG 480
QY 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer 345
Db 481 CCGCGCTTTAACTCAAGAGCGCTCCCAAGTACTGTAAGTGAAGTGGCGAGCCCTGAGC 540
QY 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365
Db 541 GCATCGTCATCTCAGCCCACTTGGTCATCTCTCGCATCTCTCGCATCTTGTGGCCATGACCTG 600
QY 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
Db 601 TTGGCCCTAACTGGACCTCGAGCGCGATGGGGGCGAGATGATGAGATCAACGGAGGAC 660
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCCAGCAGTTGGCGCTGTGCCAACCGACGCTCTCCCTATACCCC 705
```

RESULT 12

US-10-038-854-37

; Sequence 37, Application US/10038854

; Publication NO. US20040022781A1

; GENERAL INFORMATION:

; APPLICANT: Spytex, Kimberly A

; APPLICANT: Li, Li

; APPLICANT: Wolenc, Adam R

; APPLICANT: Vernet, Corine

; APPLICANT: Eisen, Andrew J

; APPLICANT: Liu, Xiaohong

; APPLICANT: Malyankar, Uriel M

; APPLICANT: Shinkets, Richard A

; APPLICANT: Tchernev, Velizar

; APPLICANT: Spaderna, Steven K

; APPLICANT: Gorman, Linda

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Patturajan, Meera

; APPLICANT: Gusev, Vladimir Y

; APPLICANT: Gangolli, Esha A

; APPLICANT: Guo, Xiaojia S

; APPLICANT: Shenoy, Suresh G

; APPLICANT: Rastelli, Luca

```
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
```

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; CURRENT APPLICATION NUMBER: US/10/038,854

; CURRENT FILING DATE: 2003-01-22

; PRIOR APPLICATION NUMBER: 60/258,928

; PRIOR FILING DATE: 2000-12-29

; PRIOR APPLICATION NUMBER: 60/259,415

; PRIOR FILING DATE: 2001-01-02

; PRIOR APPLICATION NUMBER: 60/259,785

; PRIOR FILING DATE: 2001-01-04

; PRIOR APPLICATION NUMBER: 60/269,814

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/279,832

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,833

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,863

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/283,889

; PRIOR FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: 60/284,447

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/286,683

; PRIOR FILING DATE: 2001-04-25

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 411

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 37

; LENGTH: 8645

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-038-854-37

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 1.92e-88 | Length: | 8645 |
| Score: | 975.00 | Matches: | 210 |
| Percent Similarity: | 62.41% | Conservative: | 49 |
| Best Local Similarity: | 50.60% | Mismatches: | 92 |
| Query Match: | 45.73% | Indels: | 64 |
| DB: | 17 | Gaps: | 12 |

SEQ14-X-AT-28-64-76 (1-400) x US-10-038-854-37 (1-8645)

```
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGGATGTGAAGAACGAGCGCTTACTGCTCTCCACACGAGCAGCAGAGAGGAA 210
QY 20 ArgArgTyrThrSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGGCGCTACAAATTCCTCCGACACAATGAGGAGTCCGGGTACCCACACAGAATCC 270
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTTCAGCGAGACATTGAAGCTTTTGTATCATGATTCTCGCGCTGCTTTACGCG 330
QY 58 SerArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
Db 331 AACAGATGAGAGATTTCGTTACAGAGAGACGACGAGTTCATACAGACAGGACAGAA 390
QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTACCTTAAGGCACTTAGGAGTTTGTGAACACGACCACTCGAAGAGGACTGCGATTTGT 450
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
```

| | | | | |
|--------------------------------------|------|---|------|---|
| Db | 451 | GGGAAATGGGGTCCCTCCACAGAGGTACTCTATCAGTCGAGGGTGCAGTCTGATCT | 510 | APPLICANT: Liu, Xiaohong |
| QY | 118 | GlulAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg | 137 | APPLICANT: Malyankar, Uriel M |
| Db | 511 | GAAATGAAGCAGTGATGTCCCAAGAGATGCCATGAGACTTTGGGGCAGGGGGTCAAA | 570 | APPLICANT: Shimkets, Richard A |
| QY | 138 | SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr | 157 | APPLICANT: Tchernev, Velizar |
| Db | 571 | TCAGGCGGAGCTCTGCTGCTCAAGTGGTCCAACTCAGCCCTCACCCTGACAGATACG | 630 | APPLICANT: Spaderna, Steven K |
| QY | 158 | GlulHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla | 174 | APPLICANT: Gorman, Linda |
| Db | 631 | GAGCAGGAAACAAGTCGACGAGAGAGAGTGAAGTGAACCTGCAAGCAATCAAGGCCAGTCT | 690 | APPLICANT: Kekuda, Ramesh |
| QY | 175 | ArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisAla | 194 | APPLICANT: Patturajan, Meera |
| Db | 691 | ACCTGCGAGCCCTTGGCGCCT-----TCCCATAGCAGCAGCTCTGCACAGCATCAT | 741 | APPLICANT: Gusev, Vladimir Y |
| QY | 195 | AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla | 214 | APPLICANT: Gangolli, Esha A |
| Db | 742 | CCATCCATCACTCTCTCAACAGAACTCCCTGACCAATAGAGGAACAGAGTCCGGCC | 801 | APPLICANT: Guo, Xiaojia S |
| QY | 215 | ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- | 230 | APPLICANT: Shenoy, Suresh G |
| Db | 802 | CCG-----CGGCTGCTTTGCCCGCGAGCTGCAAAACCACA | 837 | APPLICANT: Rastelli, Luca |
| QY | 231 | ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr | 249 | APPLICANT: Casman, Stacie J |
| Db | 838 | CCCGAGTCCGTCAGCTGCAGGACAGCTGGGTCCTTGGCAGTAATGTACCACTGGAAGC | 897 | APPLICANT: Boldog, Perenc |
| QY | 250 | ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnIleGluMetAsp | 269 | APPLICANT: Burgess, Catherine E |
| Db | 898 | AGG----- | 900 | APPLICANT: Edinger, Shlomit R |
| QY | 270 | IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPhelysPro | 289 | APPLICANT: Ellerman, Karen |
| Db | 901 | -----CAITTCCTATTCAAAACA | 918 | APPLICANT: Gunther, Erik |
| QY | 290 | Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer | 308 | APPLICANT: Smithson, Glennnda |
| Db | 919 | GGACAGGTACAAAGCCAGTGTTCAGTACTGCAACCCAGGATACACAAATGGCATCTGGC | 978 | APPLICANT: Millet, Isabelle |
| QY | 309 | ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe | 328 | APPLICANT: MacDougall, John R |
| Db | 979 | TCTGTATTATCACCCTACTCGGCCACTTACCTAGAAACCCCTATCAGAAAGTGTCTTT | 1038 | APPLICANT: TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same |
| QY | 329 | AsnLeuLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal | 348 | FILE REFERENCE: 21402-230 |
| Db | 1039 | AAATTCAGAGACTTCAAGTACTGTAGCTGGAATGCACCTGCTGTGCTGGCTAGGG | 1098 | CURRENT APPLICATION NUMBER: US/10/038,854 |
| QY | 349 | IleSerAlaThrLeuValIleLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu | 368 | PRIOR FILING DATE: 2003-01-22 |
| Db | 1099 | GTCTGGTCTCTGCGCAATACCTCTCTGTCTTATTTATAGCAATGCATCTCTTTGGCCCT | 1158 | PRIOR APPLICATION NUMBER: 60/258,928 |
| QY | 369 | AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp | 385 | PRIOR FILING DATE: 2000-12-29 |
| Db | 1159 | AACCTGCGACTCAGCAGACTGAAATGACATTTGAGAAATGGAAGTGAATCTGAT | 1218 | PRIOR APPLICATION NUMBER: 60/259,415 |
| QY | 386 | ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro | 400 | PRIOR FILING DATE: 2001-01-02 |
| Db | 1219 | ACC-----ATGCCAACAAACACTGTGTCTATTACCT | 1248 | PRIOR APPLICATION NUMBER: 60/259,785 |
| RESULT 13 | | | | PRIOR FILING DATE: 2001-01-04 |
| US-10-038-854-35 | | | | PRIOR APPLICATION NUMBER: 60/269,814 |
| Sequence 35, Application US/10038854 | | | | PRIOR FILING DATE: 2001-02-20 |
| Publication No. US20040022781A1 | | | | PRIOR APPLICATION NUMBER: 60/279,832 |
| GENERAL INFORMATION: | | | | PRIOR FILING DATE: 2001-03-29 |
| APPLICANT: Spytek, Kimberly A | | | | PRIOR APPLICATION NUMBER: 60/279,863 |
| APPLICANT: Li, Li | | | | PRIOR FILING DATE: 2001-03-29 |
| APPLICANT: Wolenc, Adam R | | | | PRIOR APPLICATION NUMBER: 60/283,889 |
| APPLICANT: Vernet, Corine | | | | PRIOR FILING DATE: 2001-04-13 |
| APPLICANT: Eisen, Andrew J | | | | PRIOR APPLICATION NUMBER: 60/284,447 |
| | | | | PRIOR FILING DATE: 2001-04-18 |
| | | | | PRIOR APPLICATION NUMBER: 60/286,683 |
| | | | | PRIOR FILING DATE: 2001-04-25 |
| | | | | Remaining Prior Application data removed - See File Wrapper or PALM. |
| | | | | NUMBER OF SEQ ID NOS: 411 |
| | | | | SOFTWARE: PatentIn Ver. 2.1 |
| | | | | SEQ ID NO 35 |
| | | | | LENGTH: 8675 |
| | | | | TYPE: DNA |
| | | | | ORGANISM: Homo sapiens |
| | | | | US-10-038-854-35 |
| | | | | |
| | | | | Alignment Scores: |
| | | | | Pred. No.: 1,93e-88 |
| | | | | Length: 8675 |
| | | | | Score: 210 |
| | | | | Matches: 210 |
| | | | | Percent Similarity: 62.41% |
| | | | | Conservative: 49 |
| | | | | Best Local Similarity: 50.60% |
| | | | | Mismatches: 92 |
| | | | | Query Match: 45.73% |
| | | | | Indels: 64 |
| | | | | Gaps: 12 |
| | | | | DB: 17 |
| | | | | SEQ14-X-AT-28-64-76 (1-400) x US-10-038-854-35 (1-8675) |
| QY | 1 | MetAspValIysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu | 19 | |
| Db | 151 | ATGGATGTGAAGAACGCGAGGCTTACTCTCTCTGACCAAGAGCAGCAGCAGAGAGAA | 210 | |
| QY | 20 | ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer | 38 | |
| Db | 211 | CGGCGCTACACAAATTCCTCCGACAGCAATGAGGAGTGGCGGGTACCCACACAGAAGTCC | 270 | |


```
QY 39 TyrSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TAGAGTTCCAGCGAGACATTGAAGCTTTTGGATCATGATTCCTCGCGGCTGCTTTACGGC 330

QY 58 SerArgValLysAspIle***ProGlnGluAlaGluPheCysArgThrGly***Asn 77
Db 331 AACAGAGTGAAGGATTGGTTTCCACAGAGAGGACGAGTCTCACTACAGACAGACAGAAAT 390

QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTTACCTCAAGGAGTTAGGAGTTTGAACACAGCAACTCCAGAGAGACTGCGATTGTGT 450

QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGGAATGGGCTCCCTCACAGAGGTACTCTATCAGTCAGGAGTCTGAGTGTGATCTACT 510

QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrPheGlyArgSerThrArg 137
Db 511 GAAATGAAGCAGTGTATCCCGAGAGCATGCGCATGAGACTTTGGGCGCAGGGGGTCAAA 570

QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCGGAGCTCTCGCTGTCAAGTCGGTCCAACCTCAGCCCTCAGCTGACAGATACG 630

QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCAGGAAACAAAGTCCGACAGTGAATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690

QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db 691 ACCCTGCAGCCCTTGGCGCT-----TCCATAAGCAGCACTCTGCACAGCATCAT--- 741

QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCATCACTCTCTCAACAGAAACTCCCTGACCAATAGAGAAACAGAGTCCGGCC 801

QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CGGCTGCTTTGCGCCCGCAGCTGCAAAACACA 837

QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCGCAGTCCGTCAGCTGAGGACAGCTGGTCTTGGCAGTAATGACACTGGAAAGC 897

QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 898 AGG----- 900

QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTTCTATTCAAAACA 918

QY 290 Gly----GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GCAACAGGTACAAACGCACTGTTCACTGCAACCCAGGATACAAATGCACTGCTGCG 978

QY 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTCACCACTTACTCGGCCACTTACTAGAAACACCTATCAAGAAAGTGCTTT 1038

QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAGAAGCTCTCAAGTACTGTAGCTGGAATGCACTGCACTGTGTCGCGTAGGG 1098

QY 349 IleSerAlaThrValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTCTCTCGCAATACTCTCTGCTTATTTATAGCAATGCACTCTCTTGGCCTC 1158

QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGCTACAGCACTGAAATGACATTTGGAATGGAATGGAATGGAATTTCTGAT 1218

QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
```

```
Db 1219 ACC-----ATGCCAACAAACACTGTGTGCTATTACCT 1248
```

RESULT 14

```
US-10-144-194A-81
; Sequence 81, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 9695
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (435)..(8336)
US-10-144-194A-81
```

Alignment Scores:

```
Pred. No.: 1,68e-84 Length: 9695
Score: 937.00 Matches: 206
Percent Similarity: 63.57% Conservative: 54
Best Local Similarity: 50.37% Mismatches: 93
Query Match: 43.95% Indels: 56
DB: 16 Gaps: 13
```

SEQ14-X-AT-28-64-76 (1-400) x US-10-144-194A-81 (1-9695)

```
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArg---AspAlaGlu 19
Db 435 ATGGATGTAAGAGACCGCGCA---CACCGCTCTTTGACCAGAGGAGCGCTGTGGCAAGAG 491

QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 492 TGTGCTACACAAAGCTCTCTCTGGACAGTGAAGCTGCGGCTGCCACACAGAAATCC 551

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 552 TACAGCTCCAGTGAGACTCTGAAGGCTATGACCATGACAGAGGATGCACTATGGAAC 611

QY 59 ArgValLysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPhe 78
Db 612 CAGGTACACAGACTCATCCCGGAGTCAGATGAGTTTCTTAGACAGGAACCAACTTC 671

QY 79 ThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThr 98
Db 672 ACCCTTGGCGAACTGGCATCTGTGAGCCCTCC---CCACACGAGCGGCTACTGCTCC 728

QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 729 GACATGGGATCCTTCCAGGGCTACTCCCTTAGCAGAGGCTGACGCGGACTCGGAC 788

QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTyrPheGlyArgSerThrArgSer 138
Db 789 ACCGAGGAGGATGTTCTCCAGAACCCCATCAGACTGTGGGCGAGGGATAAATCC 848

QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 849 AGGCGCAGTTCGGCTCTCCAGTCTGTGAAACTCGGCCCTTACCTGACTGACTGACTGAC 908

QY 159 HisGluAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArg-----Leu 176
Db 909 AACGAAACAAATCAGATGATGAGAACCGGCTCCGAAACCAACCAAGCGACTGCTG 968

QY 177 ArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaSer 196
Db 969 AGGCCCCCTCTCCNACCC---CCTCACAAACACACAGCTGTCCCATCACAC---TCGTCC 1022
```

```
QY 197 IleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn-----ProSer 212
Db 1023 GCCAACTCCCTCAACAGGAATCTACTACCAATCGCGGAGTCAGATCCAGCCCGCCGCC 1082
QY 213 ProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluPro 232
Db 1083 CCAGCGCCCAATGAC-----CTGGCCACCACACCA-----GAGTCC 1118
QY 233 AlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeu 252
Db 1119 GTTCAGCTTCAGGACAGCTGGCTGCTAAACAGCAACGTCGCCACTGGAGACCGG----- 1172
QY 253 GlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGly 272
Db 1172 ----- 1172
QY 273 AlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLys---ProGlyGly 291
Db 1173 -----CACITTCCTCTTCAAGACCTCTCTCGGGG 1199
QY 292 ThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyr 311
Db 1200 AGCACACCTTGTTTCAGCAGCTCTTCCCGGGGATACCTTTGACCTCAGGAACCGTTTAC 1259
QY 312 SerProProArgProGluProArgSerThrPheAlaArgProAlaPheAsnLeuLys 331
Db 1260 ACGCCCCCGCCCGCTCTGCCCGAGGAATCTTCTCCAGAGGCTTTCAAGCTGAAG 1319
QY 332 LysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSerAla 351
Db 1320 AAGCCCTCCAAATACTGACGTGGAATGTGTGCTCCCTCTCGCCATTGCCGCGCCCTC 1379
QY 352 ThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrpHis 371
Db 1380 CTCITGGCTATTCTGCTGCGGTATTTCATAGCAATGCATCTGCTCGGACTCAATTGGCAA 1439
QY 372 LeuGlnProMetGluGlyGlnMetTyrGluMetTyrGluIleThrGluAspThrAlaSerTrpPro 391
Db 1440 CTCACGCTCGAGATGGGACACCTTT-----AACATGGGATTAAGACCGGCTTACCA 1493
QY 392 ValProThrAspValSerLeuTyrPro 400
Db 1494 GGAAACGATGATGGCAACAATGCCA 1520

RESULT 15
US-09-808-602-75
; Sequence 75, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herriman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 2496
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-808-602-75
```

```
Alignment Scores:
Pred. No.: 5,05e-82 Length: 2496
Score: 905.00 Matches: 208
Percent Similarity: 54.28% Conservatives: 52
Best Local Similarity: 43.42% Mismatches: 85
Query Match: 42.45% Indels: 134
DB: 9 Gaps: 14

SEQ14-X-AT-28-64-76 (1-400) x US-09-808-602-75 (1-2496)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArg---AspAlaGlu 19
Db 1 ATGGATATAAAGATCGAAGA---CACCGCTCTTTGACGAGAGCGCGTTCGGGAAGAG 57
QY 20 ArgArgTyrThrSerSerAla**SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 58 TGTGCTATATACTATTCTTCACTCGACAGTGAAGACTGCAGAGTACCACTCAGAGTCC 117
QY 39 TyrSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 118 TACAGCTCCAGTGAGACTCTGAAAGCATATGACCATGACACGAGGATGCACCTACGGAAT 177
QY 59 ArgValLysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPhe 78
Db 178 CGAGTTTCAGACCTGGTTTCACAGGGAGTCGATGAGTTTCCAAGGCAAGGAACGAACTTC 237
QY 79 ThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThr 98
Db 238 ACCCTTGCAAGACTGGGAAATCTGTGAGCCCTCT---CCCCATCGAAGTGGCTACTGCTCG 294
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 295 GACATAGGAATACTCCATCAAGGCTATTCTTGAGCACTGGCTCTGTATGCTGACTCAGAC 354
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSer 138
Db 355 ACGAGGCGCGGATGTCTCCAGAGCACCGCATAGGCTGTGGGGAAGAGGATCAAAATCC 414
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 415 AGGCGAAGTTCTGGCTGTCAAGTCGTGAAACTCGGCTCTCACGCTCACTGACTCCGAC 474
QY 159 HisGluAsnThrGluThrAspHisProGly----- 168
Db 475 AATGAGAAACAAGTCAGATGAGGAAAAACGGTCTGCCATTCCACCTACATCCTCGTCTAGC 534
QY 168 ----- 168
Db 535 CTTCCTCCATCTGCTCAGTCGCCAGTTCTCATATTCCTCCACAGTTAGTGTGCCAGATG 594
QY 168 ----- 168
Db 595 CCATTGTAGACAGCAATACGTCCTCCATCAATCATGGACACCAATCCTGACGAGGATTC 654
QY 169 -----Gly 169
Db 655 TCTCCTAATTCATACCTACTAAGACATGTTTCAGGGCCACAGCAGGATCCAGCAGTGGC 714
QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrPro---ProProProLeuSerHis 185
Db 715 CCTTCAAAACCATCAGCAGGATCAAGCTGAGGCCACCTCTCCCTCTCTCACAACAC 774
QY 186 AlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 775 TCGCTGTCC-----CATCATCACTCGTCTGCAACTCTCCCTCAACAGGAACCTCGCTC 825
QY 206 ThrProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSer 221
Db 826 ACCAAACCGCCGCAACCAAGATCCACGCGCTGTCTCCGCTCCCAATGAC-----CTGGCG 879
QY 222 GlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeu 241
```

```
Db      880 ACCACGGCT-----GAGTCTGTGCAGCTGCAGGACAGCTGGTGCTC 921
QY      242 AsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeu 261
Db      922 AACAGCAACGTGGCTGGAGCCAGG----- 948
QY      262 GlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSer 281
Db      948 ----- 948
QY      282 AspGlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db      949 -----CATTTCTGTTTAAGACATCTTCTGGAACGACTCCGCTGTTTCAGTAGCTCTTCC 1002
QY      301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db      1003 CCTGGCTTACCCACTGACCTCAGAACAGTTTATACTCCACCTCCAGGCTGTACCTAGA 1062
QY      321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db      1063 AATACATTTTCCAGGAATGCATTCAAGCTGAAAAGCCCTCCAAGTATTGTAGCTGAAA 1122
QY      341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db      1123 TGTGCTGCTTTATCTGCAATTGCTGTGCAGTCTGCTTGCCATCTCTGTAGCATATTTC 1182
QY      361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db      1183 ATAGCGATGCACCTCTGGGGCTGAACCTGGCAGCTGAGCCCGCGACGACACACCTTC 1242
QY      381 GluIle-----ThrGluAspThrAlaSerSerTrpPro 391
Db      1243 AGCAACGGGCTGGCGCGGGCGGGCGGGAGGAGCGAGCGGGCGGCACCT 1299
```

Search completed: August 15, 2004, 00:56:30
Job time : 659 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 19:26:15 ; Search time 103 Seconds
(without alignments)

2155.149 Million cell updates/sec

Title: SEQ14-X-AT-28-64-76

Perfect score: 2132

Sequence: 1 MDVKEKPYRSLRRDAER.....EITDTASSWPVTDVSLYP 400

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xih
-Q=/cgn2_1/USPTO.spool/MITRA020/runat_06082004_114102_237/app_query.fasta.1.583
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MITRA020 @cgn2_1.1.69 @runat_06082004_114102_237 -NCPU=6 -ICPU=3 -NO_MMAPP
-LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|----|------------------|
| 1 | 2122 | 99.5 | 1680 | 3 | US-08-891-845-3 |
| 2 | 2122 | 99.5 | 1680 | 4 | US-09-514-573-3 |
| 3 | 2122 | 99.5 | 3111 | 3 | US-08-891-845-1 |
| 4 | 2122 | 99.5 | 3111 | 4 | US-09-514-573-1 |
| 5 | 1285 | 60.3 | 2387 | 3 | US-08-891-845-11 |
| 6 | 1285 | 60.3 | 2387 | 4 | US-09-514-573-11 |
| 7 | 131.5 | 6.2 | 2949 | 4 | US-09-623-326-6 |
| 8 | 131.5 | 6.2 | 4446 | 4 | US-09-231-899-69 |
| 9 | 129.5 | 6.1 | 1860 | 4 | US-09-051-994-1 |
| 10 | 128.5 | 6.0 | 8438 | 1 | US-07-945-283-1 |
| 11 | 127.5 | 6.0 | 4456 | 4 | US-09-095-443-1 |
| 12 | 126.5 | 5.9 | 4403765 | 3 | US-09-103-840A-2 |

| | | | | | | | |
|---|----|-------|-----|---------|---|----------------------|--------------------|
| c | 13 | 126.5 | 5.9 | 4411529 | 3 | US-09-103-840A-1 | Sequence 1, Appli |
| | 14 | 126 | 5.9 | 2507 | 2 | US-08-471-066B-1 | Sequence 1, Appli |
| | 15 | 125 | 5.9 | 1998 | 4 | US-09-252-991A-11748 | Sequence 11748, A |
| c | 16 | 125 | 5.9 | 3180 | 4 | US-09-252-991A-11618 | Sequence 11618, A |
| | 17 | 123.5 | 5.8 | 2499 | 4 | US-09-758-282B-100 | Sequence 100, App |
| | 18 | 123.5 | 5.8 | 2517 | 4 | US-09-758-282B-147 | Sequence 147, App |
| | 19 | 123.5 | 5.8 | 2517 | 4 | US-09-758-282B-149 | Sequence 149, App |
| | 20 | 123.5 | 5.8 | 2517 | 4 | US-09-758-282B-152 | Sequence 152, App |
| | 21 | 123.5 | 5.8 | 2517 | 4 | US-09-758-282B-168 | Sequence 168, App |
| | 22 | 123.5 | 5.8 | 2517 | 4 | US-09-758-282B-171 | Sequence 171, App |
| | 23 | 123.5 | 5.8 | 2517 | 4 | US-09-758-282B-180 | Sequence 180, App |
| | 24 | 123.5 | 5.8 | 2517 | 4 | US-09-758-282B-183 | Sequence 183, App |
| | 25 | 123.5 | 5.8 | 2517 | 4 | US-09-758-282B-186 | Sequence 186, App |
| | 26 | 123.5 | 5.8 | 2733 | 4 | US-09-623-326-1 | Sequence 1, Appli |
| | 27 | 123 | 5.8 | 4078 | 4 | US-09-016-434-1152 | Sequence 1152, App |
| | 28 | 123 | 5.8 | 4553 | 4 | US-09-023-655-975 | Sequence 975, App |
| | 29 | 123 | 5.8 | 4689 | 3 | US-09-105-537-34 | Sequence 34, Appl |
| | 30 | 123 | 5.8 | 5121 | 4 | US-09-252-991A-15189 | Sequence 15189, A |
| | 31 | 123 | 5.8 | 36778 | 3 | US-09-105-537-5 | Sequence 5, Appli |
| | 32 | 123 | 5.8 | 38506 | 4 | US-09-320-878-19 | Sequence 19, Appl |
| | 33 | 123 | 5.8 | 38506 | 4 | US-09-141-908-1 | Sequence 1, Appli |
| | 34 | 123 | 5.8 | 38506 | 4 | US-09-657-440-19 | Sequence 19, Appl |
| c | 35 | 123 | 5.8 | 71989 | 4 | US-09-443-501A-2 | Sequence 2, Appli |
| | 36 | 122.5 | 5.7 | 1686 | 2 | US-08-648-657-1 | Sequence 1, Appli |
| | 37 | 122.5 | 5.7 | 1689 | 2 | US-08-648-657-2 | Sequence 2, Appli |
| | 38 | 122.5 | 5.7 | 2517 | 4 | US-09-758-282B-174 | Sequence 174, App |
| | 39 | 122.5 | 5.7 | 2588 | 4 | US-09-480-017-3 | Sequence 3, Appli |
| | 40 | 122.5 | 5.7 | 2943 | 4 | US-09-379-530B-3 | Sequence 3, Appli |
| | 41 | 122.5 | 5.7 | 49377 | 1 | US-08-764-233A-1 | Sequence 1, Appli |
| | 42 | 121.5 | 5.7 | 1682 | 3 | US-09-096-399-1 | Sequence 1, Appli |
| | 43 | 121.5 | 5.7 | 1696 | 3 | US-09-096-399-3 | Sequence 3, Appli |
| | 44 | 121.5 | 5.7 | 1899 | 4 | US-09-640-358-3 | Sequence 3, Appli |
| | 45 | 121.5 | 5.7 | 1965 | 4 | US-09-640-358-11 | Sequence 11, Appli |

ALIGNMENTS

RESULT 1
US-08-891-845-3
; Sequence 3, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: F1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168


```
QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCGCTCGTCCGCGGACAGCGAGGAGGCAAGCGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCTGAAGCGCTACGACACAGGAGCGCGCTAGCTATGGCAGCGCGCTC 513
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 514 AAGGACATTTGTCGCGCAGAGCGCGAGAAATTCGCGCACAGGTGCCAACTTCACCTG 573
QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGGAGCTGGGCTGGAAAGATTAACGCCCTTCACGGGACCTGTACCGGACAGACAT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCGCCCTGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTCAGACACCGCGTGTGTGGGGCGGAGACACCGGTACGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTGCTGTCCAGCGGGCAATTCATCTCACACTCACCGACACCGGACATGAA 813
QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCGGGCGGCTGCAGAACACCGCGGGTCCGGACCGCGCGC 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CGCGCGCTCTCGACCGCCACACCCCAACCCAGCACCGCGGCTCCATTAACCTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGGCACTTCACGCGGAGAGCAACCCCGCGCGCCCGCCCGACGACCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCGCCCTCGCGCGCGCGCGAGGAGCTGCCACCGCCCGGAGGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCTCGAGACAGCAAACTAGGCAAGCAGCGCATTCCTAGGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGACGACCACTCATTCAGATGGACATTCGCGCGCTCCGCGCATGATGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCCTCTTCAGCGCTGGAGGACCTCCCGCGCTCTCTGACACCATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 1234 CCAGGTATCCCTGACGTCCAGCAGATGTACTCTCCCGCGCGCGCCCTGCGCGCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCGCGCGCGCTTTAACTCAAGAGCCCTCCCAAGTACTGTAACTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1354 TCGCAGCGCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCACTCTGCTGGCATATCTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCATGACCTGTTGGCTTAACCTGCACCTGCAGCGCGATGGAGGGCGAGATGAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
```

```
Db 1474 GAGATCAGGAGGACACAGCCAGAGTTGGCTGTGGCAACCGAGCTCTCCCTATACCCC 1533
RESULT 4
US-09-514-573-1
; Sequence 1, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-09-514-573-1
Alignment Scores:
Pred. No.: 5,43e-188 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x US-09-514-573-1 (1-3111)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGACGTGAAGGAGGAGGAGCTTACCGCTCGTACCGCGCGCGCGCGAGCGC 393
QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCGCTCGTCCGCGGACAGCGAGGAGGCAAGCGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCTGAAGCGCTACGACACAGGAGCGCGCTAGCTATGGCAGCGCGCTC 513
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
```


Db 514 AAGGACATTGTGCGCAGGAGCGGAGGAAATTCCTGCGCAGACAGGTGCCAACTTCAACCTG 573
Qy 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGGAGCTGGGCTGGAAGAAGTAACGCGCCCTCAACGGAGCCCTGACGGACACACATT 633
Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCGCTGCTGGGCTACTCCATGGGGCTGGCTCTGTATGCCGACATGGAGCTGAC 693
Qy 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTGAGACACCCCGTGGCTGTGTGGGCGGAGACACAGCTCAGGGGCG 753
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTCCCTGTGTCAGCGCGGCGCAATTCATCTCACACTCACCGACACCGAGCATGAA 813
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaAtrGLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCTCGGGGGCTGTCAGAACACCGCGGCTCCGGACGCGCGG 873
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGGCTCTCGCAGCGCCACACCCCAACAGCAGCACCGGGCTCCATTAACTCCCTG 933
Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTACCGCCGAGGAGCAACCCAGCGCGGCGCCCGCAGCAGCTCGCTC 993
Qy 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCCCTGCGGGGGCGCCAGAGGCTGCCACGCCCGCAGGAGAACTGGCTG 1053
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCGCTGGAGACACAGAAACCTAGGCAAGCAGCCATTCCTAGGACA 1113
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGAGAGACAACTCATGTAGATGGACATTCCTCGCGGCTCCCGCCATGATGGGCTTAC 1173
Qy 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTACGGGCACTTCCTCTTCAGCCCTGGAGGACCTCCCGGCTCTCTGCACCAATCA 1233
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CAGGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCGCGCCCGCAGCCCTCGCCGCG 1293
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACTTCGCGCGGCGGCTTTAACTCAAGAGGCGCTCCAAAGTACTGTAACTGGAAG 1353
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGCAGACCTCTGAGCGCCATCGTATCTCAGCCACTCTGTCTATCTCTGTCGACATCTT 1413
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGACCTGTTGGCTAAACTGGCCACCTGCGAGCCGATGGAGGGGCGAGATGTAT 1473
Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCAGCAGTGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1533

RESULT 5

US-08-891-845-11

; Sequence 11, Application US/08891845

; Patent No. 6098873

; GENERAL INFORMATION:

; APPLICANT: Schaefer, Gabriele M.

; APPLICANT: Sliwowski, Mark

; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-891-845-11

Alignment Scores:

Pred. No.: 4.83e-110 Length: 2387
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.27% Indels: 0
DB: 3 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-08-891-845-11 (1-2387)

Qy 166 HisProGlyGlyLeuGlnAsnHisAlaAtrGLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGGCGCTGCGAGAACCCAGCGCGGCTCCGGACGCGCGCGCGCTCTCGCAC 60
Qy 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCCACACCCCAACAGCAGCACCGCGGCTCCATTAACTCCCTGAACCGGGGCACTTC 120
Qy 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 AC0CGGAGGAGCAACCCAGCGCGGCGCCACGAGCACCCTCGCTCTCGGAGAGCCCTC 180
Qy 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
Db 181 GCGCGGCGCGCCAGGAGCGCTGCCACGCGCCAGAGAACTGGCTGTCTCAACAGCAATC 240
Qy 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
Db 241 CCCCTGAGACCAAGAAACCTAGGCAAGCAGCCATTCTTAGGGACATTGAGGCAACCTC 300
Qy 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGGACATTCTCGGCGGCTCCCGCATGATGGGGCTTACAGTACGCGGCACTTC 360
Qy 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305

Db 361 CTCCTTACAGCCTGGAGGACCTCCCGCTCTTCTGCACACATCACCAGGGTACCCACTG 420
Qy 306 ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 ACGTCAGCACAGTGTACTCTCTCGGCCCGCAGCCCTGCCCGCAGCACCTTCCGCCGG 480
Qy 326 ProAlaPheAsnLeuLysProSerLysTyrCysAsnTyrLysCysAlaAlaLeuSer 345
Db 481 CCGGCCCTTTAACTCAAGAGCCCTCCAAAGTACTGTAACTGGAAGTGCACACCTTGAGC 540
Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMethHisLeu 365
Db 541 GCCATCGTCATCTCAGCCACTCTGGTCACTCTCTGCTGGCATACTTGTGGCCATGACCTG 600
Qy 366 PheGlyLeuAsnTyrHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
Db 601 TTTGGCCTAAACTGGCACCCTGAGCGGATGGAGGGGCGAGATGATGAGATCAGCGAGGAC 660
Qy 386 ThrAlaSerSerTyrProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCCAGCAGTGGCTGTGTCACACCGACGCTCTCCCTATACCCC 705

RESULT 6

US-09-514-573-11
; Sequence 11, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwkowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-09-514-573-11
Alignment Scores:
Pred. No.: 4.83e-110 Length: 2387
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 60.27% Indels: 0
DB: 4 Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x US-09-514-573-11 (1-2387)
Qy 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGGCTCTCAGAACCCACGCGGGCTCCGAGCGCCGCCCGCTCTCGCAC 60
Qy 186 AlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCACACACCCCAACAGCACCCGCGCTCATTAATCTCTGAAACCGGGGCACTTC 120
Qy 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 AGCGCGAGGAGCAACCCAGCCCGGCCCCACGACCACTCGCTCTCCGAGAGCCCTC 180
Qy 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGlnAsnTyrLeuAsnSerAsnIle 245
Db 181 GCCGCGCGCGCCAGAGGCTGCCACGCCCGAGGAACTGGCTGCTCAACAGCAACATC 240
Qy 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
Db 241 CCCCTGGAGACCAAGAACCTAGCAAGCAGCCATTCCTTAGGGACATTGAGGACCACTC 300
Qy 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGGACATTCTCGCGCTCCCGCCCATGATGGGGCTTACAGTGACGGGCACTTC 360
Qy 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 361 CTCTTCAAGCTGGAGGACCTCCCGCTCTTCTGCACCAACATCACCGGGTACCCACTG 420
Qy 306 ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 ACGTCCAGCACAGTGTACTCTCTCCGCCCGGACCCCTGCCCGGAGCAGCCTTCGCCCG 480
Qy 326 ProAlaPheAsnLeuLysProSerLysTyrCysAsnTyrLysCysAlaAlaLeuSer 345
Db 481 CCGGCCCTTTAACTCAAGAGCCCTCCAAAGTACTGTAACTGGAAGTGCAGCCCTGAGC 540
Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMethHisLeu 365
Db 541 GCCATCGTCATCTCAGCCACTCTGGTCACTCTGTCATCTCTGGCATCTTGTGGCCATGCACTG 600
Qy 366 PheGlyLeuAsnTyrHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
Db 601 TTTGGCCTAAACTGGGACCTCTGAGCCGATGGAGGGGCGAGATGATGAGATCAGCGAGGAC 660
Qy 386 ThrAlaSerSerTyrProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCCAGCAGTGGCTGTGTCACACCGACGCTCTCCCTATACCCC 705

RESULT 7

US-09-623-326-6
; Sequence 6, Application US/09623326
; Patent No. 6607883
; GENERAL INFORMATION:
; APPLICANT: Frey et al.
; TITLE OF INVENTION: Polymerase Chimerae
; FILE REFERENCE: 4894
; CURRENT APPLICATION NUMBER: US/09/623,326
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: DE 198 10 879.6
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2949
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polynucleotide

```
US-09-623-326-6
Alignment Scores:
Pred. No.: 0.0321 Length: 2949
Score: 131.50 Matches: 101
Percent Similarity: 37.61% Conservative: 31
Best Local Similarity: 28.77% Mismatches: 115
Query Match: 6.17% Indels: 104
DB: 4 Gaps: 20

SEQ14-X-AT-28-64-76 (1-400) x US-09-623-326-6 (1-2949)
QY 3 ValysGluArgLysProTyrArg-----SerLeuThrArgArgArg 16
Db 1665 GTACGACAGGAGAACCTTGTAGAGTGGTCTTACTTAGGAAGCCCTACGAAGAAACGA 1724
QY 17 AspAlaGluArg-----ArgTyrThrSerSerAla***SerGluGluGly 32
Db 1725 AGTAGCTCCAAACAGCAAGTGAAGAGGAGTATCAAGAGGCTCAGGAGAGCTACAC 1784
QY 33 -----LysAlaProGlnLys-SerTyrSerSerGluThrLeuLysAlaTyrAs 49
Db 1785 AGGTGGATTCTGCGCCCTGGACGTGGCTATCTCAGGCGCTTGTCCCTGGAGTGGCGGA 1844
QY 49 pGlnAspAlaArgLeuAlaTyrGlySerArgValLysAspIle***ProGlnGluAlaG 69
Db 1845 GGAGATCCCGCCCTC-----CTCTTT---GACGAGTGGGCTTCCCGCCATCGCAAGAC 1868
QY 69 uGluPheCysArgThrGly***AsnPheThrLeuArgGluLeuGly---LeuGluGluVa 88
Db 1869 GGTCTTCCGCTGCGCGCCACCCCTCAACCTCACTCCCGGACAGCTGGAAGGTT 1928
QY 88 lThrProHisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCysGlyTyrSe 108
Db 1929 C-----CTCTTT---GACGAGTGGGCTTCCCGCCATCGCAAGAC 1967
QY 108 rMetGlyAlaGly-----SerAspAlaAspMetGluAlaAspThrValLeuSe 124
Db 1968 GGAAGACCGGCAAGCGCTCCACGCGCGCTCTGAGGCC-----CTCCG 2018
QY 124 rProGluHisProValArgLeu-TrpGlyArgSerThrArgSerGlyArgSerSerCysL 144
Db 2019 CGAGGCCACCC-----ATCTGTGAGAAGATCTCGAGTACCGGAGCTACCA--- 2068
QY 144 euSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThrGlu 164
Db 2069 --AGCTGAAGAGCACCT-----ACATTGACCCCTTCCCGGACCTCATCCACC--- 2113
QY 164 hrAspHisProGly-GlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeu 183
Db 2114 -----CCAGGACGG-----CCGCTTCCA--- 2132
QY 184 SerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGly 203
Db 2133 ---CACCCGCTTCAACACAGACGCGCACCGGCGAGCTAAGTAGCTC----- 2180
QY 204 AsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGlu 223
Db 2181 -----CGATCCCAACCTCCAGAACATCCCGTCCGACCCCGCTTGGGCAGAG 2228
QY 224 ProProAlaGlyGlyAlaGlnGluProAlaHis-AlaGlnGluAsnTrpLeuLeuAsnSe 243
Db 2229 GATCCCGCGGC-----CTTCATCGCGAGAGGGTGGTATTGTGGC 2273
QY 243 rAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlnAs 263
Db 2274 C-----CTGGACTA 2282
QY 263 pAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspG 283
Db 2283 TAGCCAGATAGACTCAGGTGTGGCC-----CACCTCCGGCGACGAGAACCT 2333
QY 283 yHisPheLeuPhe-LysProGlyGlyThrSer-----ProLeuPheCysThr 2399
```

```
Db 2334 GATCCGGTCTTCCAGGAGGCGGCGGACATCCACACGAGACCGGAGCTGGATGTCGG 2393
QY 299 hrSerProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuP 319
Db 2394 CGTCCCGGAGGCGCGTGGACCCCTGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 2453
QY 319 roArgSerThrPheAlaArgPro 326
Db 2454 GGTCTCTACGGCATGTGCGGCC 2476

RESULT 8
US-09-231-899-69
; Sequence 69, Application US/09231899
; Patent No. 6566583
; GENERAL INFORMATION:
; APPLICANT: Lassner, Michael
; APPLICANT: Metz, James G
; APPLICANT: Facciotti, Daniel
; TITLE OF INVENTION: SCHIZOCHYTRIUM PKS GENES
; FILE REFERENCE: CGNE.131.02US
; CURRENT APPLICATION NUMBER: US/09/231,899
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; EARLIER APPLICATION NUMBER: 09/090,793
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 4446
; TYPE: DNA
; ORGANISM: Schizochytrium aggregatum
US-09-231-899-69

Alignment Scores:
Pred. No.: 0.0614 Length: 4446
Score: 131.50 Matches: 103
Percent Similarity: 32.72% Conservative: 39
Best Local Similarity: 23.73% Mismatches: 167
Query Match: 6.17% Indels: 125
DB: 4 Gaps: 19

SEQ14-X-AT-28-64-76 (1-400) x US-09-231-899-69 (1-4446)
QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 3037 CGCGAGTGGCGAGCTCTATTGCTGGATTGAAGCGCTCGCGGCAAGCCATCTACTCG 3096
QY 41 SerSerGluThrLeuLysAlaTyrAsp-----GlnAspAlaArgLeuAla 55
Db 3097 TCGTGGCAGTGTAACCTCTGCGCGGACGTGGCCAGCGCGTGGCGATGCCGATGCCAG 3156
QY 56 TyrGlySerArgValLysAspIle----- 63
Db 3157 CTCGGTGGCGGCTCTCGGCGCATCGTTCATGCTCGGCGGTCTCGCGGACCGCTCATC 3216
QY 64 -----***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db 3217 GAGAAGAAGCTCCCGACGAGTTCGACGCGCTC-----TTTGGCACC 3258
QY 81 ArgGluLeuGlyLeuGluVal-ThrProHisGlyThrLeuTyrArgThrAspI 100
Db 3259 AAGGTCAACCGTCTCGAGAACCTCTCTCGCGCGC-----TCGAC 3297
QY 100 eGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAla 120
Db 3298 CGGCGCAACTCAGCACATGGTCTCTTACGCTCGCTCGCGGCTTCCACGCGCAACGCTC 3357
QY 120 pThrValLeuSer-----ProGluHisProValArgLeuTrpGly- 133
Db 3358 GGCCAGTCTGACTACCGCATGGCCACGAGGCGCTTAAACAGATGGCGCTCGAGCTCGCC 3417
```

```

QY 134 -ArgSerThrArgSerGlyArgSerCysLeuSer-SerArgAlaAsnSerAsnLeuT 153
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3418 AAGGACGTCCTCGTCAAGTCATCTGCTCGGTCCTGGACGGTGGCATGTGACGGCG 3477
QY 153 hrLeuThrArgThrGluHisGluAsnThrGluThrAspHisPro-----167
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3478 CAGCTCAAGAAGCAGTTCAGGAGATGGCGTGCAGATCATCCCGCGAGGGCGCGCT 3537
QY 168 -----GlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerH 185
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3538 GATACCGTGGCGC-----CATCGTGTCTCGCTCTCGCGGCTGAGATCCTTGTGCGC 3591
QY 185 IsAlaHisThrProAsnGln-----HisHisAlaAla-----195
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3592 AACTGGCGGACCCCTCCAGAGGTGCGCTCGGCACCATCACCTGCACCGCAGATT 3651
QY 196 -----SerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerProA 214
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3652 TCCGCCAAGTCCAAACCCCTTCTCGAGGA-----CCAGT---CATCCAGGCGCGC 3699
QY 214 laProThrAspHis-----218
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3700 CCGGTGCTGCCCATGACGCTGCCCATTTGGCTCGCTCGCGGACCTGCGCTCTTTC 3759
QY 218 -----218
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3760 CCGGGTACTCGCTCGGGCCATTGACGAGCGCCAGCTCTTCAAGGGTGTCACTGTCGAC 3819
QY 219 -----SerLeuSerGlyGluPro---ProAlaGlyGlyAla-----GlnG 231
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3820 GCGGACGTCAACTGCGAGGTGACCTTCAACCCGCTGACGCGCGCTCGGGCGCGTCAAC 3879
QY 231 luProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArg- 250
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3880 GTCCAGGCCACGCTCAAGACCTTTTCCAGCGGCAAGCTGTGTCGCGCCCTACCGCGCCGT 3939
QY 251 -----AsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIle 266
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3940 ATCGTGTCTTCAACAGCGCGCGCGCCCGGCCAACCGCACCATGACGCGCGCTCGCTC 3999
QY 267 GluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTySerAspGlyHisPheLeu 286
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4000 GATCGCGAT-----CCGGCGCTCCAGGGCTCGTCTACGACGCGCAAGACCCCTC 4047
QY 287 PheLys---ProGly-GlyThrSerProLeuPheCysThrThrSerProGlyTyProLe 305
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4048 TTCCAGGCGCGCGCTTCCGCGGCATCGATGACGTCTCTCGTGCCACCAAGCCAGCTT 4107
QY 305 uThrSerSerThrValTySerProProProArgPro-----LeuProArgSe 321
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4108 GTGGCAAGTGCAGCGCTGTCCTCCGCTCCGACGCGCTCGCGCGAGTTTGCACGGAC 4167
QY 321 rThrPheAlaArgProAlaPheAsnLeuLysProSerLysTySerAsnTrpLysCy 341
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4168 ACTGACGCCCATGACCCCTTCTGTAACGACCTGCGCTTTTCAAGGCCATGCTCGTGGTG 4227
QY 341 sAlaAlaLeuSerAlaIleValIleSerAlaThr 352
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4228 CGCCGACGCTCGGCGAGGTGCGCTCCCAACT 4261

```

RESULT 9

```

US-09-051-994-1/c
; Sequence 1, Application US/09051994A
; Patent No. 6602683
; GENERAL INFORMATION:
; APPLICANT: REGENERON PHARMACEUTICALS, INC.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE EPH FAMILY LIGANDS
; FILE REFERENCE: REG-341-PCT-US
; CURRENT APPLICATION NUMBER: US/09/051,994A
; CURRENT FILING DATE: 1998-04-24
; EARLIER APPLICATION NUMBER: PCT/US96/17201
; EARLIER FILING DATE: 1996-10-25
; EARLIER APPLICATION NUMBER: 60/007,015

```

```

; EARLIER FILING DATE: 1995-10-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (202)..(1224)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1717)
; OTHER INFORMATION: n=a, c, g, or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1741)
; OTHER INFORMATION: n=a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1782)
; OTHER INFORMATION: n=a, c, g, or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1795)
; OTHER INFORMATION: n=a, c, g, or t
US-09-051-994-1

```

Alignment Scores:

```

Pred. No.: 0.0238 Length: 1860
Score: 129.50 Matches: 111
Percent Similarity: 32.35% Conservative: 42
Best Local Similarity: 23.47% Mismatches: 177
Query Match: 6.07% Indels: 144
DB: 24 Gaps: 24

```

SEQ14-X-AT-28-64-76 (1-400) x US-09-051-994-1 (1-1860)

```

QY 1 MetAspValLysGluArgLysProTyArgSerLeuThrArgArgAspAlaGluArg 20
Db 1648 CTAGAGACGGAGAGAGGGAA-----AGATAAGACATAGAGAAGAG 1607
QY 21 ArgTyThrSerSerSera***SerGluGluGlyLysAlaProGlnLysSerTySer 40
Db 1606 AAA-----ACCAGTCAGAGGAGAGAAAGCGGTGCCAACCAAAAGGTGGCTGT 1556
QY 41 SerSerGluThrLeu-----LysAlaTyAspGlnAspAlaArgLeuAlaTyGlySer 58
Db 1555 TCCCGCCCTAGCTCTGCGTGGTGAGGGGACAGGATGCTTGGGTCTATGTCAGAGAG 1496
QY 59 ArgValLysAspIle---***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
Db 1495 GAGGCTGACCTGTGCCCTCCCCAAGGAA-----AAAGGATCCAGG 1454
QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyArg 97
Db 1453 GACACTGACAGACAGGGGGG-----ACCATGGCCAAACGGGAGG 1412
QY 98 ThrAspIleGly-----LeuProGlnCys 105
Db 1411 CAGGAAGTGGGCGAGTGGGAATCCTTAAGGAATCCTTAAGTGGAGAGGAGGAGCGCC 1352
QY 106 GlyTySerMetGlyAlaGlySerAspAlaAspMetGlu---AlaAspThrValLeuSer 124
Db 1351 GGGAGGAGTGAAGGGGCTGGGGGGGCACAGGGGGCCGAGATGTAGAGGTGTCCTCAAA 1292
QY 125 Pro-----GluHisProValArgLeuTyr-----GlyArg 134
Db 1291 CCAGGAATTAAACTGGAGGAGGACCCCAAGAGGGGCTGGATTTCAGGATAGCCACGTGAGA 1232
QY 135 SerThrArgSerGlyArgSerCysLeuSerSerArgAlaAsnSerAsn---LeuThr 153
Db 1231 GGAGCCCTCATACCTTGTAGTAGAAGTTTGGAGGGGCTTGGGGGGGCTCTGGGGGGG 1172

```



```
QY 63 e***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeuArgGluLe 83
Db 5829 CTGCGGGGACGGGGGGGGCCCTGCAGACAGACCC-----5790
QY 83 uGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIleGlyLeuPr 103
Db 5789 -----GCCGGCGCGCGCGGA-GAGGGCGCCCTCGCGCGCGCG-----5749
QY 103 oGlnCysGlyTyrSerMetGlyAla-----GlySerAspAlaAspMetGlu---AlaAs 120
Db 5748 -----GGCTTCAGCAGCTCAGCAGCGCGCGCGCAGCAGCTCGGACCTCTCCCGCGCG 5696
QY 120 pThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyAr 140
Db 5695 CTGCGCCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGGCTC 5636
QY 140 gSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGl 160
Db 5635 CTCCTCG-----TCCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5582
QY 160 uAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPr 180
Db 5581 ACAACAGCAGGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5528
QY 180 oProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLe 200
Db 5527 GTCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5480
QY 200 uAsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLe 220
Db 5479 CTCGCCGACCTCTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5434
QY 220 uSerGlyGluPro--ProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrp 239
Db 5433 -----GAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5381
QY 240 LeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGly 259
Db 5380 GCTATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5324
QY 260 ThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAla 279
Db 5323 GCCTCTGGCG-----5315
QY 280 TyrSerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThr 299
Db 5314 -----CGGCTCGACCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5274
QY 300 SerProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuPro 319
Db 5273 GCGCGCGGACTCCCGCGAGGGGCTCTGGACGAGGACGACGTCGCGCGCGCGCGCGCGCG 5214
QY 320 ArgSerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrp 339
Db 5213 ---GTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5166
QY 340 -----LysCysAlaAlaLeuSerAla 346
Db 5165 GGGACTCGAGGAGCAGCAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5118
```

RESULT 11

US-09-095-443-1

; Sequence 1, Application US/09095443

; Patent No. 6342593

; GENERAL INFORMATION:

; APPLICANT: Plowman, Gregory

; APPLICANT: Peles, Eior

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT

; TITLE OF INVENTION: OF ALP RELATED DISORDERS

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

```
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,443
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/049,477
FILING DATE: June 12, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 235/055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4456 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-095-443-1
Alignment Scores:
Pred. No.: 0.146 Length: 4456
Score: 127.50 Matches: 93
Percent Similarity: 32.9% Conservative: 37
Best Local Similarity: 23.6% Mismatches: 162
Query Match: 5.98% Indels: 103
DB: 4 Gaps: 16
SEQ14-X-AT-28-64-76 (1-400) x US-09-095-443-1 (1-4456)
QY 2 AspValLysGluArgLysProTyrArgSer-----LeuThrArgArg 15
Db 1357 GAGCTGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1416
QY 16 ArgAspAlaGluArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro 35
Db 1417 GAGGAGAGTGTAG-----GCAGTGAAGAGCAGAGAGAGAGAGAGAGAGAGAGAG 1452
QY 36 GlnLysSerTyrSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 55
Db 1453 -----GAGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1485
QY 56 TyrGlySerArgValLysAspIle***Pro-----65
Db 1486 GCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1544
QY 66 GlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeuArgGluLeuGlyLeu 85
Db 1545 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1604
QY 86 GluGluValThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCys 105
Db 1605 TGGTACCTACTCGGGCGCGCGCG-----1625
QY 106 GlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrValLeuSerPro 125
Db 1626 -----CCAGCTGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1649
QY 126 GluHisProValArgLeuTrpGlyArgSerThrArgSerGly-----ArgSerSerCys 143
```

```

; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 9.92e+03 Length: 4403765
Score: 126.50 Matches: 79
Percent Similarity: 36.42% Conservatives: 43
Best Local Similarity: 23.58% Mismatches: 121
Query Match: 5.93% Indels: 92
DB: 3 Gaps: 16

SEQ14-X-AT-28-64-76 (1-400) x US-09-103-840A-2 (1-4403765)
QY 32 GlyLysAlaProGlnLysSerTyrSerSerGluThrLeuLysAlaTyrAspGlnAsp 51
Db 1334300 GTCGCGGTCCCAATCTGCATCAGGTGCTCAAAATCTCTTCAC----- 1334256
QY 52 AlaArgLeuAlaTyrGlySerArgValLysAspIle***ProGlnGluAlaGluGluPhe 71
Db 1334255 -----CACGGCCCAAGCCACGAGAT-----GACTTT 1334229
QY 72 CysArgThrGly***AsnPheThrLeuArgGluLeuGluGluValThrProPro 91
Db 1334228 CGCGGGCAAGTGAACAACTTCGTCGGGAGCCGGGATACGACGTGT-----CCC 1334175
QY 92 HisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAla 111
Db 1334174 CGTGGG-----AGAACGGGA-----CGATGTGGTCTGTCTCCGTGCC 1334136
QY 112 GlySer-AspAlaAspMetGluAlaAspThrValLeuSerProGluHisProVal-ArgL 131
Db 1334135 ACCAGCGGACGGGAGCTTTCAC-----CTCGTCAGCGGATCCCGACGTCCGGG 1334085
QY 131 euTrpGlyArgSerThrArgSerGlyArgSerCysLeuSer-----SerArgA 148
Db 1334084 CAAGCGCATCATCGCGGACGGCGACGGCGCGGCGCGCGCGCGCGCGCGCGCG 1334025
QY 148 laAsnSerLeuLeuThrLeu-ThrAspThrGluHisGluAsnThr----- 162
Db 1334024 CGTCGAGGAACATCGCTTGAATCTCGGCGGACGCCAGCAATCCGTCGCGCGCGGTG 1333965
QY 163 -----GluThrAspHisProGlyGlyLeu 170
Db 1333964 AGACGAGCCCATACAGTCGAGGGCAGGGACGCGACGGCGCGCGCGCGCGCGCG 1333905
QY 171 GlnAsnHisAlaArgLeuArgThrProProProLeuSerHisAlaHisThrProAsn 190
Db 1333904 AGCTCGCACCCAGCGCGGGGTGCGGCCACCTGACGAGCGGCGCGCGCGCGCGAAC 1333845
QY 191 Gln-HisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAs 210
Db 1333844 CAAGCGCATCAAAACCGCGCTAATCGGTCGCGCGCGCGCGCGT-----CGGCGCGACG 1333791
QY 210 nProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyAlaGl 230
Db 1333790 CGCCGAGGACACCGCGGGGACCAACCGGTC-----GGCGCGCACG 1333740
QY 230 nGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrAr 250
Db 1333739 CCAGGGTGTATGGGCCCCCGCGCGAGCGGCCACACGCGCATCTT---GTGATGCCGA 1333683
QY 250 gAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspI 270
Db 1333682 GCGTGTGCGCGATGGTCCG-----CAGAT 1333659
QY 270 euLeuGlyAlaSerArgHisAspGly-AlaTyrSerAspGlyHisPheLeuPheLysPro 290
Db 1333658 CGTCGCGC---GAACGCCAAGATGGTTTCTGACTGATCGCGCGCTCGAGGCGCGCATCCGG 1333602
QY 290 lycGlyThrSerProLeu----- 295
Db 1333601 GCGCATCGAGCCCATCACAGCAATATTGTGTGTCGGGTAGACCCGGGTTCGTGTCG 1333542
```

```

Db 1650 AGGCCCCATGCAAT-----GCGGTAGACACTGGGCTCGCTCTTACCCAGCCCTGTC 1703
QY 144 LeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThrGlu 163
Db 1704 CTACACACGGAGCTGGGCTTGTGCGCGATCTCCCTCCACAGCATGGCGGT---GAG 1760
QY 164 ThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeu 183
Db 1761 CAGTCCCTATGTGGGGGTAGGCGCGGCCACCACTGTCAGAGTCTCCCTCGCGGCCACC 1820
QY 184 Ser-----HisAlaHisThrProAsnGlnHisHisAlaAla 195
Db 1821 TCCTCAATCTCAGGCCCGGATGGCCATCGCGTTCGGCCAGCCACACACAGTAGA 1880
QY 196 -SerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAlaPr 215
Db 1881 TAGCATCAGCGGCCCATCCCGACCCACACAGCCCGCCAGCGGCAAAACCCACCCCTGCTCC 1940
QY 215 oThrAspHisSerLeuSerGlyGluProPro-----Al 226
Db 1941 TCCCCCGCTGCTTCCTGTCGCCCCCAGCGACGCACTGCCCCAGCTTACACCTACCC 2000
QY 226 aGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrp-LeuLeuAsnSerAsnIleP 246
Db 2001 TGCAGGGGCTAAGCAACCATCCCGACAGCAGCACCATTCTCTTCTGGATCCCGACAGG 2060
QY 246 roLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuI 266
Db 2061 TTTTCCAGCCCCAGGATTTGGGCCCCAGCCCC-----AGCCCCATCTCTCAGCCCCCATCC 2114
QY 266 leGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr---SerAspGlyHisP 285
Db 2115 TTCACAGCGT---TTGGGCTCAGCCCCCAGCAGCAGCCCTTCCACTCCAGCATCCACA 2171
QY 285 he-LeuPheLysProGlyGlyThrSerProLeu-----PheCys--- 297
Db 2172 TCTCTTCCGACCCAGGCCCCAGGACTCTTACCCCAACATCCCTTACCCCTATGCCCC 2231
QY 298 -----ThrThrSerProGlyTyrProLeuThrSerSerThr 309
Db 2232 TCAGCTGGGGTCTGTGGGCGAGCGCGCCACCCCTACACACCGCTTACCCAGGTCC 2291
QY 310 ValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsn 329
Db 2292 CGCTAAGACCTCTGCGACGCCACTCAGGGGCTCTGCTTTCGCCAGCCCTGGGCCCC 2351
QY 330 LeuLysLysProSerLysTyrCysAsnTrpLysCys 341
Db 2352 TCAGCTCTCCCATCCCCCACTGGCATATGTCCTGTC 2387
```

RESULT 12

```

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
```


ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.
 REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: FORS-01800
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2507 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-471-066B-1

Alignment Scores:
 Pred. No.: 0.0808 Length: 2507
 Score: 126.00 Matches: 98
 Percent Similarity: 37.13% Conservative: 29
 Best Local Similarity: 28.65% Mismatches: 123
 Query Match: 5.91% Indels: 92
 DB: 2 Gaps: 19

SEQ14-X-AT-28-64-76 (1-400) x US-08-471-066B-1 (1-2507)

QY 10 ArgSerLeuThrArgArgAspAlaGluArgGlyTyrSerSerSerAla***Ser 29
 Db 1273 AGGCTCTTTGGCTTTACCGGAGGTGAGAGGCCCTTTCCGCTGCTCGCCACATG 1332
 QY 30 Glu---GluGlyLysAlaProGlnLysSerTyrSerSerGluThrLeuLysAlaTyr 48
 Db 1333 GAGCCACGGGGGCGCCCTGGAGCTGCCTATCTCAGGGCTTGTCCCTGGAGGTGCC 1392
 QY 49 AspGlnAspAlaArgLeuAlaTyrGlySerArgValLysAspIle***ProGlnGluAla 68
 Db 1393 GAGGAGATCGCCGCTC---GAGGCC 1416
 QY 69 GluGluPheCysArgThrGly***AsnPheThrLeuArgGluLeuGly---LeuGluGlu 87
 Db 1417 GAGTCTTCGGCTGGCGGCCACCCCTTCAACCTCAACTCCCGGGACCAAGCTGAAAGG 1476
 QY 88 ValThrProHisGlyThrLeuThrArgThrAspIleGlyLeuProGlnCysGlyTyr 107
 Db 1477 GTC-----CTCTTT---GACGAGCTAGGCTTCCCGCATCGGCAAG 1515
 QY 108 SerMetGlyAlaGly-----SerAspAlaAspMetGluAlaAspThrValLeu 123
 Db 1516 ACGGAGAGACCGGCAAGCGCTCCACCGCGCGCTCCTCTGGAGGCC-----CTC 1566
 QY 124 SerProGluHisProValArgLeu-TripGlyArgSerThrArgSerGlyArgSerSerCy 143
 Db 1567 CGCGAGCCCAACCC-----ATCGTGAGAGATCTCTGAGTACCGGGAGCTCACCA-- 1618
 QY 143 sLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThrG1 163
 Db 1619 ---AGCTGAAGAGCACT-----ACATTGACCCCTTGGCGGACCTATCCACC-- 1663
 QY 163 urThrAspHisProGly-GlyLeuGlnAsnHisAlaArgLeuArgThrProProProL 183
 Db 1664 -----CCAGGACGGG-----CCGCTTCCA- 1682
 QY 183 euSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArg 203
 Db 1683 -----CACCCGCTTCAACAGACGGCCACGGCCACGGCGAGGCTAAGTAGCTC----- 1730
 QY 203 lyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyG 223
 Db 1731 -----CGATCCCAACCTCCAGAACATCCCGCTCCGACCCCGCTTGGGAG 1776
 QY 223 luProProAlaGlyGlyAlaGlnGluProAlaHis--AlaGlnGluAsnTrpLeuLeuAsn 242
 Db 1777 AGGATCCCGCGGC-----CTTCATCGCGGAGGGGTGGTATTGGTG 1821

QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln 262
 Db 1822 GCC-----CTGGAC 1830
 QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
 Db 1831 TATAGCCAGATAGAGCTCAGGGTGTGTGCC-----CACCTCTCCGGCGACGAGAAC 1881
 QY 283 GlyHisPheLeuPhe-LysProGlyGlyThrSer-----ProLeuPheCysTh 298
 Db 1882 CTGATCCGGTCTTCCAGAGGGCGGAGCATCCACACGGAGACCGCCAGCTGGATGTC 1941
 QY 298 rThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLe 318
 Db 1942 GCGTCCCGCGGAGCGCTGACCCCTGATGCGCGCGGCGGCGGACCATCACTTC 2001
 QY 318 uProArgSerThrPheAlaArgProAlaPheAsnLeuLys---ProSerLysTyr 336
 Db 2002 GGGGTCTCTACGGCATGTGCGGCCACCGCTCTCCAGAGAGTAGCCATCCCTTAC 2059
 RESULT 15
 US-09-252-991A-11748
 ; Sequence 11748, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 11748
 ; LENGTH: 1998
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-11748

Alignment Scores:

Pred. No.: 0.07 Length: 1998
 Score: 125.00 Matches: 85
 Percent Similarity: 33.42% Conservative: 40
 Best Local Similarity: 22.73% Mismatches: 128
 Query Match: 5.86% Indels: 122
 DB: 4 Gaps: 17

SEQ14-X-AT-28-64-76 (1-400) x US-09-252-991A-11748 (1-1998)

QY 5 GluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArgGlyThrSer 24
 Db 716 CACGCGCACAGCTGCAAGCCCTGGAGCGCGCAACGACGCTCGCTCGCGGACGAGTTC 775
 QY 25 SerSerAla***SerGluGlyLysAlaProGlnLysSerTyrSerSerSerGluThr 44
 Db 776 GAGCGCGCAACGAACCTCCCGGGCTTCCGGCGCCAGGCCAACACGCTCGCGGAAC 835
 QY 45 -----LeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgValLys 61
 Db 836 CGCGCGCAGCTCAGCGCGGTGAGCAGGAC-----CTGGCGAAAGCGAGCCAGCC 889
 QY 62 AspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeuArg 81
 Db 890 ACCTTGCTGTGTCACCGCGCGCGAGAGCGGCATCGCCACCGCC-----GTGCTCGCC 940
 QY 82 GluLeuGly-----LeuGluGluValThr 89
 Db 941 GAAGCGGGCAGACCGTCGACAGCTCGCGTCCGCTGCTGAGCATCGTTCCCGCGGACAC 1000

```

Qy 90 ProProHisGlyThrLeuTyr-----ArgThrAspIleGlyLeuProGlnCysGlyTyr 107
Db 1001 CCGTTGACGCCGAACTCTACGCCCGCAGCAAGTCCATCGGTTCATCCCGCGCGC--- 1057
Qy 108 SerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrValLeu-SerProGluHi 127
Db 1058 -----GACCGGTGCTGATCCGCTACCAG 1081
Qy 127 sProValArgLeuTrpGlyArgSerThrArgSerGlyArgSerSerCysLeuSerSerAr 147
Db 1082 GCCTATCCGTACC-----AGAAATTGCGCCAGTACCACGCGCAGGTGCAGTCGATCTCC 1135
Qy 147 gAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThrGluThrAsp----- 165
Db 1136 CCGCGC--CAGCGTCTCTCTATGCCGAGCTTTCAGCATGTCGCGCGCGGTACCGGGGCTCG 1193
Qy 166 -----HisProGlyGly----- 169
Db 1194 GCCAGATGGCGAGCAGCTGTACCGCTCGGGTAACCTCGACGACCAGCGCGGTGACCG 1253
Qy 170 -----LeuGlnAsnHisAlaArgLeuArgThrProProPro-- 181
Db 1254 CCTACGGCCAGCGCGTCCGCTGCAGAGCGGCATGCTGTGACGCGCACATCCTCCAGG 1313
Qy 182 -----ProLeu-----SerHisAlaHisThrProAsnGlnHisAl 194
Db 1314 ACACCCGCGCCTCTACGATGGTGTGTGAACCGCTCTACAGCCTGACCGGCAAACTCT 1373
Qy 194 aAlaSerIleAsnSerLeuAsnArgGlyAsnPhetThrProArgSerAsnProSerProAl 214
Db 1374 AGGAACGCCCATGGCTTTCTCGA-----CGCTCTCGCCCTGCGCCTGG 1418
Qy 214 aProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHi 234
Db 1419 GCCG-----CCGCTGCGCTGGTGTGTCAGACCGAAGCCA 1454
Qy 234 sAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLy 254
Db 1455 CCGA-----ATCGCGCTGGCCTGCGCTGGCGATGATCGCGGTACCAGC 1499
Qy 254 sGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSe 274
Db 1500 GCCACCA----- 1506
Qy 274 rArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysProGlyThrSerPr 294
Db 1507 -----TACCGGCTGATGAACTGCGCC 1529
Qy 294 oLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProPr 314
Db 1530 GCGCTTCTCCGTATCGCTCAAGGGCATCTCCCTCAAGCACTGATCCAGACGCCCCACC 1589
Qy 314 oProArgPro---LeuProArgSerThrPheAlaArgPro 326
Db 1590 GCCTCGGCTGGGTACCGCGCGGTGAA---GCTCGACCT 1626

```

Search completed: August 15, 2004, 02:14:11
Job time : 5435 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 18:05:13 ; Search time 3336 Seconds
(without alignments)
3580.597 Million cell updates/sec

Title: SEQ14-X-AT-28-64-76
Perfect score: 2132
Sequence: 1 MDVKERKPYRSLRRRAER.....EITEDTASSWPFDVSLYP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPIO.spool/MITRA020/runat_06082004_114101_218/app_query.fasta_1.583
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MITRA020 @CGN_1_1_3437@runat_06082004_114101_218 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_vrt:*
22: em_gss_vrt:*
23: em_gss_mam:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| C 1 | 924 | 43.3 | 620 | 14 | CB532245 |
| C 2 | 924 | 43.3 | 620 | 14 | CB532612 |
| C 3 | 884 | 41.5 | 566 | 13 | BU072782 |
| C 4 | 769 | 36.1 | 625 | 13 | BU200362 |
| C 5 | 725 | 34.0 | 878 | 13 | BU474934 |
| C 6 | 720.5 | 33.8 | 3447 | 11 | AK047423 |
| C 7 | 699 | 32.8 | 884 | 13 | BQ735622 |
| C 8 | 685 | 32.1 | 402 | 14 | CB544750 |
| C 9 | 680 | 31.9 | 789 | 13 | BU474512 |
| C 10 | 675.5 | 31.7 | 887 | 13 | BU119163 |
| C 11 | 657 | 30.8 | 4556 | 11 | AK034286 |
| C 12 | 583.5 | 27.4 | 490 | 9 | AL922332 |
| C 13 | 568 | 26.6 | 870 | 10 | BG036207 |
| C 14 | 519.5 | 24.4 | 511 | 29 | AY405291 |
| C 15 | 519.5 | 24.4 | 511 | 29 | AY405292 |
| C 16 | 517.5 | 24.3 | 511 | 29 | AY405293 |
| C 17 | 512.5 | 24.0 | 534 | 10 | BF944080 |
| C 18 | 509 | 23.9 | 2716 | 11 | AK031198 |
| C 19 | 490.5 | 23.0 | 880 | 13 | BX760712 |
| C 20 | 477.5 | 22.4 | 588 | 10 | BF953105 |
| C 21 | 470 | 22.0 | 2627 | 11 | AK050784 |
| C 22 | 456.5 | 21.4 | 643 | 10 | BB657984 |
| C 23 | 448 | 21.0 | 717 | 10 | BB654584 |
| C 24 | 438.5 | 20.6 | 593 | 29 | CE756467 |
| C 25 | 433 | 20.3 | 284 | 10 | BF388169 |
| C 26 | 412 | 19.3 | 259 | 28 | CC178145 |
| C 27 | 397 | 18.6 | 825 | 13 | BU253342 |
| C 28 | 388.5 | 18.2 | 864 | 14 | CF225019 |
| C 29 | 369 | 17.3 | 864 | 13 | BU172491 |
| C 30 | 362 | 17.0 | 397 | 13 | BY010856 |
| C 31 | 362 | 17.0 | 704 | 13 | BY733523 |
| C 32 | 355 | 16.7 | 643 | 13 | BY723994 |
| C 33 | 331.5 | 15.5 | 608 | 9 | AL955817 |
| C 34 | 330.5 | 15.5 | 707 | 13 | BU229930 |
| C 35 | 329.5 | 15.5 | 801 | 13 | BU451304 |
| C 36 | 287 | 13.5 | 461 | 13 | BU473055 |
| C 37 | 282.5 | 13.3 | 751 | 13 | BU461850 |
| C 38 | 273.5 | 12.8 | 1060 | 29 | CNS04977 |
| C 39 | 272.5 | 12.8 | 745 | 29 | CF434866 |
| C 40 | 269.5 | 12.6 | 404 | 29 | CG514853 |
| C 41 | 257.5 | 12.1 | 386 | 9 | AA702693 |
| C 42 | 256.5 | 12.0 | 674 | 28 | B2111628 |
| C 43 | 248 | 11.6 | 474 | 9 | AI753786 |
| C 44 | 245 | 11.5 | 1013 | 12 | BM017187 |
| C 45 | 237 | 11.1 | 1125 | 28 | CC310899 |

ALIGNMENTS

| RESULT 1 | LOCUS | CB532245/c | 756697 MARC 6BOV Bos taurus cdna 3', mRNA sequence. | 620 bp | mRNA | linear | EST 16-MAY-2003 |
|---|------------|---|---|--------|------|--------|-----------------|
| CB532245/c | DEFINITION | 756697 MARC 6BOV Bos taurus cdna 3', mRNA sequence. | | | | | |
| CB532245 | ACCESSION | CB532245 | | | | | |
| CB532245.1 | VERSION | CB532245.1 | GI:29395750 | | | | |
| EST | KEYWORDS | EST | | | | | |
| Bos taurus (cow) | SOURCE | Bos taurus (cow) | | | | | |
| Bos taurus | ORGANISM | Bos taurus | | | | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos. | | | | | | | |
| 1 (bases 1 to 620) | REFERENCE | 1 | | | | | |

81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
 289 CGCGAGTGGGCTGGGGAGGTACACCCCGCAGCGGACGCTGTACCGCAGCAGATC 348
 101 GlyLeuProGlnCysGlyTyrSerMetGlyValaGlySerAspAlaAspMetGluAlaAsp 120
 349 GGCTGCGCCCACTGGCGCTACTCCTCGGCGCCAGCTCTGAGGCGGAGCTGGAGCGCGAC 408
 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
 409 CGCGCACTGTCGCCCGAGCACCCCGTGGCGCTGTGGCGCCGACGACCGGTGAGGACGC 468
 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
 469 AGTCTCTGCTGTCACCGCGGGCAACTCAACCTCAGCTCACTGACACGCGGACGAG 528
 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
 529 AACACGGAGACCGATCACCCGGCGGCTCGAGAACCACTTCTGACTCCCGACACCGCCG 588
 181 ProProLeuSerHisAlaHisThrPro 189
 589 CCTCGCTCTCGCATGCCACACCCCC 615

BU072782 566 bp mRNA linear EST 27-AUG-2002
 im47g02.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6038211 5'
 similar to TR:014667 O14667 PRO-NEUREGULIN-1, GAMMA ISOFORM
 PRECURSOR ;, mRNA sequence.
 BU072782
 BU072782.1 GI:22513971
 EST.

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 566)
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
 Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
 Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R.,
 Williams, T., Jackson, Y., and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmeldon@biochp.harvard.edu
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@im.wustl.edu)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 339.
 Location/Qualifiers
 1. 566
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6038211"
 /tissue_type="Purified pancreatic islet"
 /lab_host="PH108"
 /clone_lib="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site 1:
 NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
 Size-selected on agarose gel. Average insert size ~1kb. 5'

XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."

ORIGIN

Alignment Scores:

Pred. No.: 8.16e-55 Length: 566
 Score: 884.00 Matches: 162
 Percent Similarity: 95.86% Conservative: 0
 Best Local Similarity: 95.86% Mismatches: 7
 Query Match: 41.46% Indels: 0
 DB: 13 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BU072782 (1-566)

QY 165 AspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSer 184
 Db 57 GATCATCCGGCGGCTCGAGAACCAACCGCGGCTCGGACCGCGCGCGCTCTCG 116
 QY 185 HisAlaHisThrProAsnGlnHisHisAlaAlaSerlleAsnSerLeuAsnArgGlyAsn 204
 Db 117 CAGCCCAACACCCCAACACGACCAACCGCGGCTCTCAATTAACCTCCCTGAACCGGGGCAAC 176
 QY 205 PheThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluPro 224
 Db 177 TTCACGCGGAGGAGCAACCCAGCGCGCCCGACGACCACTCGCTCTCGGAGAGCCC 236
 QY 225 ProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsn 244
 Db 237 CCTGCGCGCGCGCGCGGAGGCTGCGCCACGCGGAGAACTGGCTGCTCAACAGCAAC 296
 QY 245 IleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsn 264
 Db 297 ATCCCCCTGGAGACCAAGAAACCTAGGCAAGACGACCACTTCTAGGACATTTGAGGCAAC 356
 QY 265 LeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHis 284
 Db 357 CTCATTGAGATGACATTTCTCGCGGCTCCCGCCATGATGGGCTTACAGTACGCGGCAC 416
 QY 285 PheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrPro 304
 Db 417 TTCTCTTCAAGCTGGAGGACCTCCCGCTCTTCTGACCAACACATCACCGAGGTACCCA 476
 QY 305 LeuThrSerSerThrValTyrSerProProProProArgProLeuProArgSerThrPheAla 324
 Db 477 CTGACGTCGACGACAGTGACTCTCCCTCCGCGGACCCCTGCGCGGACCACTTTTCG 536
 QY 325 ArgProAlaPheAsnLeuLysPro 333
 Db 537 CGGGCGGCTTTTAACTCAAGAGCCT 563

RESULT 4

BU072782

LOCUS

DEFINITION

604158648F1 CSEQCHN03 Gallus gallus cDNA clone ChEST1010p11 5',
 mRNA sequence.

BU072782

BU072782.1 GI:25363728

EST.

KEYWORDS

SOURCE

ORGANISM

Gallus gallus (chicken)

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

1 (bases 1 to 625)

Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 22335534

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

FEATURES

source

PUBMED 12445392
 CONTACT: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source
 1. 625
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CHEST1010p11"
 /tissue type="whole embryo"
 /dev stage="20-21"
 /lab_host="DH10B"
 /clone_lib="CSEQOCHN03"
 /note="Organ: whole embryo; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN
 Alignment Scores:
 Pred. No.: 2,426-46 Length: 625
 Score: 769.00 Matches: 149
 Percent Similarity: 89.94% Conservative: 12
 Best Local Similarity: 83.24% Mismatches: 16
 Query Match: 36.07% Indels: 2
 DB: 13 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BU200362 (1-625)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
 DB 93 ATGGATGTAAGAAAGAAACCGTATCGATCTCTGACTCGGCGCGGACACGGAGCGC 152
 QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
 DB 153 CGCTACACAGCTCTTCAGCCGAGAGTGGAGGAGCGAGCGCTCTCAGAGGCTCTATAGC 212
 QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
 DB 213 TCCAGTGAGACCTGAGGCTTATGATCAAGATCCAGTTCAGCTACAGCAATCGGTC 272
 QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPhetThrIleu 80
 DB 273 AAAGACATGGTCACCGAGGCTGATGAATTCGCGGAGGAGCGCAACTCTCTTTG 332
 QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
 DB 333 CGGAGCTGGCTTGAAGATGTGACTCCACCATGGGACTTTGTACCGGACTGATAT 392
 QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
 DB 393 GGGTGCTCTCACTGTGGCTACTCCATCAGCGGTGGCTCAGATGCCGACACGGAGCGGAT 452
 QY 121 ThrValLeuSerProGluHis-ProValArgLeuTyrGlyArgSerThrArgSerGlyVal 140
 DB 453 GTGGTCATGTACCTGAGCATCCCTGTGAGGCTCTGGGAGCGGACCAACCAATCCCGAGC 512

QY 140 gSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisG1 160
 DB 513 CAGTCTCTGCTTGTGCGGCGGCACTCAACTCACCCTCACCACGAGCAGCA 572
 QY 160 uAsn-ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArg 177
 DB 573 GAATTACGAACTGATCATCTCCAAATCTTCAAAATCAATCAAGACTCGCA 625

RESULT 5
 BU474934
 LOCUS 603364850F1 CSEQBN21 Gallus gallus cDNA clone CHEST262c14 5', mRNA
 DEFINITION sequence.
 ACCESSION BU474934
 VERSION BU474934.1 GI:25968511
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 878)
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 22335634
 12445392
 COMMENT
 Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source
 1. 878
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Layer"
 /db_xref="taxon:9031"
 /clone="CHEST262c14"
 /sex="Female"
 /dev stage="adult"
 /lab_host="DH10B"
 /clone_lib="CSEQBN21"
 /note="Organ: ovary; Vector: pBluescript II KS(+); Site 1:
 EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN
 Alignment Scores:
 Pred. No.: 6,516-43 Length: 878
 Score: 725.00 Matches: 139
 Percent Similarity: 90.91% Conservative: 11
 Best Local Similarity: 84.24% Mismatches: 14
 Query Match: 34.01% Indels: 1
 DB: 13 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BU474934 (1-878)

| | | | |
|---|----------|--|------|
| NDGKMEQVFVLTITAIKGLPCLLPFTV" | | | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 9.89e-42 | Length: | 3447 |
| Score: | 720.50 | Matches: | 169 |
| Percent Similarity: | 52.77% | Conservative: | 50 |
| Best Local Similarity: | 40.72% | Mismatches: | 101 |
| Query Match: | 33.79% | Indels: | 95 |
| DB: | 11 | Gaps: | 15 |
| SEQ14-X-AT-28-64-76 (1-400) x AK047423 (1-3447) | | | |
| Qy | 1 | MetAspValLysGluArgLysProTyrArgSerLeuThrArg--ArgArgAspAlaGlu | 19 |
| Db | 862 | ATGGAGCAACAGAGATGCAAACTTATCAGGCTCTGTCACAAAGTCAGGATCAAGATGAT | 921 |
| Qy | 20 | ArgArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyr | 39 |
| Db | 922 | CTAGCTTATACAGTCTTCTGATGAGAGTGAAGATGGAGAAACCAAGACAGTCAATC | 981 |
| Qy | 40 | SerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer--- | 58 |
| Db | 982 | AATCCAGGAAACTCTGATGAGTATACCAAGAGCTGAGGAGGATTAACAATAGCCAG | 1041 |
| Qy | 59 | -----ArgValLysAspIle***ProGlnGluAlaGlu--GluPheCysArgThrGly | 75 |
| Db | 1042 | AGTAAAGAGGAGAGATGGAGAACTACTCAAGAGATAGAATCTGTGAA----- | 1095 |
| Qy | 76 | ***AsnPheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeu | 95 |
| Db | 1096 | -----ACACCTCCT-----ACTTTG | 1110 |
| Qy | 96 | -----TyrArgThrAspIle---GlyLeuProGlnCysGlyTyrSerMetGlyAla | 111 |
| Db | 1111 | TGCTCGGCTACACACAGATGCACAGTGTCTTCGACATGGCTACCAGCTGGAGATG | 1170 |
| Qy | 112 | GlySerAspAlaAspMetGluAlaAspThrValLeuSerProGluHisProValArgLeu | 131 |
| Db | 1171 | GGATCTGATGTAGACAGACAGAGAGCTGCATCACCTGACATGCATGCATCAGATG | 1230 |
| Qy | 132 | TpGlyArgSerThrArgSerGlyArgSerCysLeuSerSerArgAlaAsnSerAsn | 151 |
| Db | 1231 | TGGATAAGGGATGAATACAGACACAGTCTCTGTCTGTCAGTAGGCGCAACTTGCA | 1290 |
| Qy | 152 | LeuThrLeuThrAspThrGluHisGluAsnThrGluThrAspHisProGlyGlyLeuGln | 171 |
| Db | 1291 | CTGTCTTGACCGACACTGATCATGAACGG---AAGTCTGATGGGAAATGGTTTAAA | 1347 |
| Qy | 171 | ----- | 171 |
| Db | 1348 | TTCTCTCTGTTGTTGTGACATGGAGGCTCCAGCTGATTCGGCTCAAGACATGCAAGC | 1407 |
| Qy | 172 | -----AsnHisAlaArgLeuArg-----ThrProProProLeuSerHisAla | 186 |
| Db | 1408 | AGCCACACACACAGTTCACCTTCAGACCCCTCCACACCCGACCTCTCCACATGTC | 1467 |
| Qy | 187 | HisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr | 206 |
| Db | 1468 | TGCACCTGTGCCAGGAAGCCACTCTCTACAGTGAGCTCTCTACAAAGAGATCAATGAT | 1527 |
| Qy | 207 | ProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProProAla | 226 |
| Db | 1528 | ACCGCAGCAGCCAGCCAGCCAGCT-----GCTCTGCTCTCTCCA | 1566 |
| Qy | 227 | GlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIlePro | 246 |
| Db | 1567 | ACCAGCACACAGGATTCAGTTCACTCTGCATTAACAGCTGGGTCTTGAACATACATCCA | 1626 |
| Qy | 247 | LeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIle | 266 |
| Db | 1627 | CTGGAGACACAGG----- | 1638 |

| | | | |
|--|------|--|------|
| Qy | 267 | GluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeu | 286 |
| Db | 1639 | -----CAITTCCTG | 1647 |
| Qy | 287 | PhelysProGly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu | 305 |
| Db | 1648 | TTCAACATGATGATCTGTTCTTCTGCTATCTTCTAGTCGACGAGTCAGACTACCTCTG | 1707 |
| Qy | 306 | ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg | 325 |
| Db | 1708 | ACATCTAATACTGCTACTTACACCCACCCAGCGCTGCTCGAGACACTTTCCCGA | 1767 |
| Qy | 326 | ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer | 345 |
| Db | 1768 | CCTGCCITCACTTTTACAAACCATACAGATGCTCAATTCGAAGTCACAGCTTGAGC | 1827 |
| Qy | 346 | AlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPheValAlaMetHisLeu | 365 |
| Db | 1828 | GCCACTGCAATCAACAGTCACTTTGGCCTTGTACTAGCCTAT--GTAATTGTACACTTG | 1884 |
| Qy | 366 | PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr | 380 |
| Db | 1885 | TTTGCCCTGACGTGGCAGTTGCACACACTT---GGACAGATCTAT | 1926 |
| RESULT 7 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| ACCESSION | | | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| COMMENT | | | |
| FEATURES | | | |
| source | | | |
| 1. .884 | | | |
| /organism="Xenopus laevis" | | | |
| /mol_type="mRNA" | | | |
| /db_xref="taxon:8355" | | | |
| /clone="IMAGE:5542635" | | | |
| /dev_stage="embryo, stage 31-32" | | | |
| /lab_host="DH10B (phage-resistant)" | | | |
| /clone_lib="NICHD_XGC_Emb4" | | | |
| /note="Organ: whole embryo; Vector: pCMV-SPORT6; Site 1: | | | |
| Not1; Site 2: SalI; Cloned unidirectionally. Primer: Oligo | | | |
| dr. Average insert size 2.1 kb. Constructed by Life | | | |
| Technologies. Note: This is a Xenopus Gene Collection | | | |
| (XGC) library." | | | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | | | |
| Score: | | | |
| Percent Similarity: | | | |
| Length: | | | |
| Matches: | | | |
| Conservative: | | | |

| | | | |
|------------|-----|---|-----|
| | 81 | ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuYrAfgInrASpile | 100 |
| QY | | ::: ::: ::: ::: ::: ::: ::: | |
| Db | 496 | CGGGAGCTGGTCTTGAAAGATGTGACTCCACCCATGGGACTTTGTACCGGCATGATAATT | 555 |
| | 101 | GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaaspMetGluAlaasp | 120 |
| QY | | ::: ::: ::: ::: ::: ::: ::: | |
| Db | 556 | GGGCTGCCTCACTGTGGTACTCCATCAGCGCTGGCTCAGATCCGACACGGAGCGCGAT | 615 |
| | 121 | ThrValLeuSerProGluHisProValArgLeuTipGlyArgSerThrArgSerGlyArg | 140 |
| QY | | ::: ::: ::: ::: ::: ::: ::: | |
| Db | 616 | GTGGTCATGTCACCTGAGCATCTCTGTGAGGCTCTGGGAGCGCACACAATCCGGACGC | 675 |
| | 141 | SerSer-CysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGI | 160 |
| QY | | ::: ::: ::: ::: ::: ::: ::: | |
| Db | 676 | AGTCTCCTGCTGTGCGAGCGGGC-AACTCCAACCTCACCTC-ACCAGACGGAGCAGA | 733 |
| | 160 | uAsnThrGluThr 164 | |
| QY | | ::: ::: ::: ::: ::: ::: ::: | |
| Db | 734 | GAATACCGAACT 746 | |
| | | | |
| RESULT 10 | | | |
| BUI19163 | | | |
| LOCUS | | 887 bp mRNA linear EST 25-NOV-2002 | |
| DEFINITION | | 603143046f1 CSEQCHL16 Gallus gallus cdna clone CHEST137111 5', mRNa | |
| | | sequence. | |
| ACCESSION | | BUI19163 | |
| VERSION | | BUI19163.1 GI:25328007 | |
| KEYWORDS | | EST, | |
| SOURCE | | Gallus gallus (chicken) | |
| ORGANISM | | Gallus gallus | |
| | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| | | Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; | |
| | | Phasianinae; Gallus. | |
| REFERENCE | | 1 (bases 1 to 887) | |
| AUTHORS | | Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., | |
| | | Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J. | |
| TITLE | | A Comprehensive Collection of Chicken cDNAs | |
| JOURNAL | | Curr. Biol. 12 (22), 1965-1969 (2002) | |
| MEDLINE | | 22335534 | |
| PUBMED | | 12445392 | |
| COMMENT | | Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST) PO Box 88, Manchester, M60 1QD, UK Tel: 01612008930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk. | |
| FEATURES | | Location/Qualifiers | |
| source | | 1..887 | |
| | | /organism="Gallus gallus" | |
| | | /mol_type="mRNA" | |
| | | /strain="Compton Line 151" | |
| | | /db_xref="taxon:9031" | |
| | | /clone="CHEST137111" | |
| | | /sex="Female" | |
| | | /tissue type="not cerebrum or cerebellum" | |
| | | /dev stage="adult" | |
| | | /lab_host="DH10B" | |
| | | /clone_lib="CSEQCHL16" | |
| | | /note="Organ: brain; Vector: pBluescript II KS(+); Site 1 | |
| | | EcoRI; Site 2: NotI; Modification of pBluescript II KS(+) | |
| | | [stratagen] vector to accommodate cDNA produced with the | |
| | | T-trimmed protocol (Construction of uni-directionally | |
| | | cloned cDNA libraries from messenger RNA for improved 3' | |
| | | end DNA sequencing by Glenn Fu et al. U.S. Patent # | |
| | | 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. | |
| | | Ligate in double stranded adaptor containing BsgI and | |
| | | BamHI sites [5'ggcgcggtgcagcccgatccgaataaaag] | |
| | | [5'aattcttttttcggatccggggtgcacg!]" | |
| ORIGIN | | | |

```
Pred. No.: 2,75e-39 Length: 887
Score: 675.50 Matches: 144
Percent Similarity: 85.79% Conservative: 13
Best Local Similarity: 78.69% Mismatches: 24
Query Match: 31.68% Indels: 5
DB: 13 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BULL19163 (1-887)

Qy 1 MetAspVallysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 343 ATGGATGTAAGAAGAAACCGTATCGATCTCTGACTCGCGCGCGACACGAGCGC 402

Qy 21 ArgTyrThrSerSerAla***SerClucluglyllysAlaProGlnLysSerTyrSer 40
Db 403 CGCTACACCAAGCTCTTCAGCGGAGGTGAGCAGCAGGCTCTCTCAGAACTCTATAGC 462

Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 463 TCCAGTCAGACCCCTGAAGGCTTATGATCAGACTCCAGGTTGACCTACAGCAATCGGGTC 522

Qy 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 523 AAAGACATGTGTGCACGAGGCTGATGAATCTCTCCGAGCAGGACCACTTCTCTTGT 582

Qy 81 ArgGlnLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 583 CGGAGCTGGTCTTGAAGATGTGACTCCCAACCCCTGGGACTTTGTACCGGACTGATATT 642

Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyValaGlySerAspAlaAspMetGluAlaAsp 120
Db 643 GGGCTGCTCACTGTGCTACTCTCATCAGCGCTGGCTCAGATCCGACACGAGGCGGAT 702

Qy 121 ThrValLeuSer-ProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 703 GTGCTCATGTCACCTGAGCATCTCTGTGAGCTCTTGGGAGCGCAACACCAATCCGAGC 762

Qy 140 gSerSerCysLeuSerSerArgAlaSerAsnLeuThrLeuThrAspThrGluHisG1 160
Db 763 CAGTCTCTGTTGTGAGCGCGGCGCACTCCAACTC-ACCTCACCAGACACGGAAGCCCG 821

Qy 160 uAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArg-ThrProp 180
Db 822 AGAATACGAAT-GATCATCTCCATCTCTCA-ATCATTCAGACTCCGAATCCACAC 879

Qy 180 roPro 181
Db 880 CTCCA 884

RESULT 11
AK034286
LOCUS
DEFINITION
Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:930172M18 product:odd Oz/ten-m homolog 1
(Drosophila), full insert sequence.
ACCESSION
AK034286
VERSION
1 GI:26083876
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
```

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4556)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. .4556
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9330172M18"
/db_xref="MGI:2398485"
/db_xref="taxon:10090"
/clone="9330172M18"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. .4556
/note="odd Oz/ten-m homolog 1 (Drosophila)
(MGI:1345185, GB|NM_011855, evidence: BLASTN, 99%,
match=3037)"

FEATURES
source

1. .4556
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9330172M18"
/db_xref="MGI:2398485"
/db_xref="taxon:10090"
/clone="9330172M18"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. .4556
/note="odd Oz/ten-m homolog 1 (Drosophila)
(MGI:1345185, GB|NM_011855, evidence: BLASTN, 99%,
match=3037)"

misc_feature

1. .4556
/note="odd Oz/ten-m homolog 1 (Drosophila)
(MGI:1345185, GB|NM_011855, evidence: BLASTN, 99%,
match=3037)"


```

SEQ14-X-AT-28-64-76 (1-400) x AL9222332 (1-490)
QY 3 VallysGluAUGlyProThrArgSerLeuThrArgArgAspAlaGluArgTyr 22
Db 489 GTGAAGGAACGAGACCCCTACCGCTCTCTGACCTCCAGCGGACACGCGCGCTAC 430
QY 23 ThrSerSerAla**SerGluGluGlyAlaProGlnLysSerTyrSerSer 42
Db 429 ACCAGCTCATCTGCGCAGCAGGAGATGCAAGATCAACCTTAAATCTTACAGCTCGAGT 370
QY 43 GluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgValLysAsp 62
Db 369 GAGACCTCAAAGCCTTCGACCAGGACTCCAGACTGGCTTATGGCAGCGGCTCAAAGAC 310
QY 63 Ile**ProGlnGluAlaGluPheCysArgThrGly**AsnPheThrLeuArgGlu 82
Db 309 CTGTGTCAACATGAGCGCGAGTTCAGACAGCAAGCGCGGACTTTCTCTCAGAGAC 250
QY 83 LeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeu 102
Db 249 ATGGCCTTCGAGATCCGTCGCCCACTATGCGAGCGTACCGGACAGAAATCGGCGCT 190
QY 103 ProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrVal 122
Db 189 CCCCACCGCGACTACTCAGTGAGTGTGTCATCAGACGCTGACACAGAAACAGCGCAT 130
QY 123 LeuSerProGluHisProValArgLeuTrpGlyArgSer---ThrArgSerGlyArgSer 141
Db 129 ATGTCCCGCCGACGCGCGCTCGCGCTCTGGGCGCGCAGCAACCAATCCGCGCGAGT 70
QY 142 SerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsn 161
Db 69 TCTGCTGTCCACAGAGCGCACTCCACCTTACGCTCCCGACCGGAGCATGAAGAC 10
QY 162 ThrGlu 163
Db 9 ACTGAG 4

RESULT 13
BG036207
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

602326960F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4428351 5',
mRNA sequence.
BG036207.1 GI:12431132
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Incyte Genomics, Inc.
Found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10178 row: h column: 16
High quality sequence stop: 713.
Location/Qualifiers
1. .870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4428351"
/tissue_type="adenocarcinoma, cell line"

FEATURES
source

```

/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_91"

/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 1.95e-31 Length: 870
Score: 568.00 Matches: 109
Percent Similarity: 99.09% Conservativeness: 0
Best Local Similarity: 99.09% Mismatches: 1
Query Match: 26.64% Indels: 1
DB: 10 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BG036207 (1-870)

```

QY 291 GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrVal 310
Db 2 GGCACCTCCCGCTCTTCTGACCATCACCAGG-TACCCACCTGACGTCACGACAGTG 60
QY 311 TyrSerProProArgProLeuProLeuProArgSerThrPheAlaArgProAlaPheAsnLeu 330
Db 61 TACTCTCTCCGCGCCGACCCCTGCGCGAGCACCTTGGCGCGCGCTTTAACTC 120
QY 331 LysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaAlaValLysSer 350
Db 121 AAGAGCCCTCCAAAGTACTGTAACTGGAAGTGGCGAGCCCTGAGCGCATCTCATCTCA 180
QY 351 AlaThrLeuValLleLeuLeuAlaTyrPheValAlaMethHisLeuPheGlyLeuAsnTrp 370
Db 181 GCCACTCTGTCTATCTCTGCTGGCATCTTTGTGGCCATGACCTGTGTGGCTTAACTGG 240
QY 371 HisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSerSerTrp 390
Db 241 CACCTGAGCGCGATGGAGGGGCGAGATGATGAGATCACGAGGACACAGCCAGGTTGG 300
QY 391 ProValProThrAspValSerLeuTyrPro 400
Db 301 CCTGTGCCAACCGACGCTCCCTATATACCCC 330

```

AY405291 511 bp DNA linear GSS 12-DEC-2003
Homo sapiens HCM2175 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

AY405291
AY405291.1 GI:39761265
GSS.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 511)

Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,

Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,

Adams, M.D. and Cargill, M.

Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios

Science 302 (5652), 1960-1963 (2003)

14671302

2 (bases 1 to 511)

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,

Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,

Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,

Adams, M.D. and Cargill, M.

Direct Submission

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

This sequence was made by sequencing genomic exons and ordering

COMMENT

them based on alignment.

FEATURES
source
Location/Qualifiers
1. .511
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>511
/locus_tag="HCM2175"

gene

ORIGIN

Alignment Scores:
Pred. No.: 3,19e-28 Length: 511
Score: 519.50 Matches: 106
Percent Similarity: 78.66% Conservative: 23
Best Local Similarity: 64.63% Mismatches: 32
Query Match: 24.37% Indels: 3
DB: 29 Gaps: 3

SEQ14-X-AT-28-64-76 (1-400) x AY405291 (1-511)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 1 ATGGATGTGAAGAAACGCGGCTTACTGCTCCCTGACCAAGCAGCAGCAGAGAGAA 60

QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 61 CGCGCGCTACACAAATTCCTCCGACAAATGAGGAGTGCGGGGTACCCACACAGAGATCC 120

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 121 TACAGTTCACAGGAGACATTGAAGCTTTTGATCATGATTCCTCGCGGTGCTTTACGGC 180

QY 58 SerArgValLysAspIle***ProGlnGluAlaGluPheCysArgThrGly***Asn 77
Db 181 AACAGAGTGAAGATTTGGTTTCCACAGAGAAGCAGACGAGTTCTACTAGACAGACAGAA 240

QY 78 PheThrLeuArgGluLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 97
Db 241 TTACCCCTAAGGCAGTAGAGTTTGTGAACCAACCACTCGAAGAGAGCTGGCATTTTGT 300

QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 301 GCGGAATGGGCTCCCTCACAGAGGTACTCTATCAGTCAGCGGTACAGATGCTGATACT 360

QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 361 GAAATGAAGCAGTGTGTTTCCACAGAGAAGCAGACGAGTTCTACTAGTCAGCGGTGCTGATACT 420

QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 421 TCAGGCGGCGAGCTCTCGCTGTCAAGTCGGTCCAACTCAGCCCTCACCCTGACAGATACG 480

QY 158 GluHisGluAsn 161
Db 481 GAGCAGGAAAC 492

RESULT 15
AY405292
LOCUS
DEFINITION
Pan troglodytes HCM2175 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION
AY405292
VERSION
AY405292.1
KEYWORDS
GSS.
ORGANISM
Pan troglodytes (chimpanzee)
Pan troglodytes
Pan troglodytes
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 511)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.

Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 511)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
Location/Qualifiers
1. .511
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>511
/locus_tag="HCM2175"

gene

ORIGIN

Alignment Scores:
Pred. No.: 3,19e-28 Length: 511
Score: 519.50 Matches: 106
Percent Similarity: 78.66% Conservative: 23
Best Local Similarity: 64.63% Mismatches: 32
Query Match: 24.37% Indels: 3
DB: 29 Gaps: 3

SEQ14-X-AT-28-64-76 (1-400) x AY405292 (1-511)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 1 ATGGATGTGAAGAAACGCGGCTTACTGCTCCCTGACCAAGCAGCAGCAGAGAGAA 60

QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 61 CGCGCGCTACACAAATTCCTCCGACAAATGAGGAGTGCGGGGTACCCACACAGAGATCC 120

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 121 TACAGTTCACAGGAGACATTGAAGCTTTTGATCATGATTCCTCGCGGTGCTTTACGGC 180

QY 58 SerArgValLysAspIle***ProGlnGluAlaGluPheCysArgThrGly***Asn 77
Db 181 AACAGAGTGAAGATTTGGTTTCCACAGAGAAGCAGACGAGTTCTACTAGACAGACAGAA 240

QY 78 PheThrLeuArgGluLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 97
Db 241 TTACCCCTAAGGCAGTAGAGTTTGTGAACCAACCACTCGAAGAGAGCTGGCATTTTGT 300

QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 301 GCGGAATGGGCTCCCTCACAGAGGTACTCTATCAGTCAGCGGTACAGATGCTGATACT 360

QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 361 GAAATGAAGCAGTGTGTTTCCACAGAGAAGCAGACGAGTTCTACTAGTCAGCGGTGCTGATACT 420

QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 421 TCAGGCGGCGAGCTCTCGCTGTCAAGTCGGTCCAACTCAGCCCTCACCCTGACAGATACG 480

QY 158 GluHisGluAsn 161
Db 481 GAGCAGGAAAC 492

Search completed: August 15, 2004, 00:43:42
Job time : 3356 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 152.319 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14 COPY_2650_2725

Perfect score: 391

Sequence: 1 SQINTVNGRRTRYDIQLQ.....EQRLREGEGGLRAWTEGK 76

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112218_29331/app.query.fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020@cgn_1.1.2156 @runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQURY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PTCUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match Length | ID | Description |
|------------|-------|--------------|----|-------------|
|------------|-------|--------------|----|-------------|

ALIGNMENTS

RESULT 1

US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

```

; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 3,166-47 Length: 8354
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-383-201-43 (1-8354)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 7982 TCCAGATCACACAGTACTTAATGCGAGGACTAGACGCTACACAGATCCAGCTCCAG 8041

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluAlaArg 40
Db 8042 TACGGGGCACTGCTGTGAACACACGCTAGCGGCAACGTTGGATGAGGAGGACCG 8101

QY 41 ValLeuGluLeuAlaArgGlnAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 8102 GTCTGGAGTGGCGGCGAGAGCGCTGCGGCAACGTTGGATGAGGAGGACCG 8161

QY 61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8162 CTGCGGGAAGGGAGGAGGCGCTGCGGCGCTGCGAGAGGGGGGAGAAG 8209

RESULT 2
US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US2004003397A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/289,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-029-020-13

Alignment Scores:
Pred. No.: 3,166-47 Length: 8354
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-029-020-13 (1-8354)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 7982 TCCAGATCACACAGTACTTAATGCGAGGACTAGACGCTACACAGATCCAGCTCCAG 8041

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluAlaArg 40
Db 8042 TACGGGGCACTGCTGTGAACACACGCTAGCGGCAACGTTGGATGAGGAGGACCG 8101

QY 41 ValLeuGluLeuAlaArgGlnAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 8102 GTCTGGAGTGGCGGCGAGAGCGCTGCGGCAACGTTGGATGAGGAGGACCG 8161

QY 61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8162 CTGCGGGAAGGGAGGAGGCGCTGCGGCGCTGCGAGAGGGGGGAGAAG 8209

RESULT 3
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

```



```
Alignment Scores:
Pred. No.: 3,16e-47 Length: 8355
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-383-201-55 (1-8355)

Qy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 7966 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGATCCAGCTCCAG 8025

Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluLysAlaArg 40
Db 8026 TACGGGCACTGTGCTTGAACACAGCTACGGGCAACGTTGGATGAGGAGGACGCG 8085

Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 8086 GTCTGGAGCTGGCCGGCAGAGAGCGTGGCCAGCGTGGCCGGCAGAGAGAGAG 8145

Qy 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8146 CTGCGGGAAGGGGAGGAGGCGCTGGGCGCTGGACAGAGGGGAGAGAG 8193

RESULT 4
US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Zhytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876

; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1

Alignment Scores:
Pred. No.: 3,2e-47 Length: 8438
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-042-865-1 (1-8438)

Qy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 8033 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGATCCAGCTCCAG 8092

Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluLysAlaArg 40
Db 8093 TACGGGCACTGTGCTTGAACACACAGCTACGGGCAACGTTGGATGAGGAGGACGCG 8152

Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 8153 GTCTGGAGCTGGCCGGCAGAGAGCGTGGCCAGCGTGGCCGGCAGAGAGAGAG 8212

Qy 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8213 CTGCGGGAAGGGGAGGAGGCGCTGGGCGCTGGACAGAGGGGAGAGAG 8260

RESULT 5
US-10-295-027-927
; Sequence 927, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
```

```

; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 927
; LENGTH: 12880
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-927

Alignment Scores:
Pred. No.: 1,35e-34 Length: 12880
Score: 308.00 Matches: 56
Percent Similarity: 89.47% Conservative: 12
Best Local Similarity: 73.68% Mismatches: 8
Query Match: 78.77% Indels: 0
DB: 16 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-295-027-927 (1-12880)

Qy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 7880 TCCGACATGACTTCTCTGTGTGAATGGGAGGACTAGACGGTTTGCAGATATTCAGCTCCAG 7939
Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 7940 CATGGAGCCCTGTGCTTCAACATCCGGTATGGGACAACTCTCGAAGAGGAAAGAATCAC 7999
Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 8000 GTGTGTGAGATTGCCACACACGCGCAGTGGCCCGAGCTGGACTAAGGAACAAGAAGG 8059
Qy 61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8060 CTGCAACAGGCGGAAGAGGGATTAGGCGCATGGATGGACAGAAGGGGAAAAAG 8107

```

RESULT 6
US-10-342-887-1743
; Sequence 1743, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1743
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1743

Alignment Scores:
Pred. No.: 3614
Score: 296.00
Percent Similarity: 86.84%
Matches: 54
Conservative: 12

```

Best Local Similarity: 71.05%      Mismatches: 10
Query Match:          75.70%      Indels: 0
DB:                   13          Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-342-887-1743 (1-3614)

QY      1  SerGlnIleAsnThrValLeuAsnGlvArgThrArgATyTrThrAspIleGlnLeuGln 20
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2924  TCGAGATCCACCACCGGTGTGAACGCGAGGACGCCAGCTTCGCGGACGTGGAGATGAG 2983

QY      21  TyrGlyValaLeuCysLeuAsnThrArgTyTrGlyThrThrLeuAspGluGluLysAlaArg 40
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2984  TTCGGCGCGCTGGCGTGCACCTGCGCTAGCGCATGACCTCGACGAGAGAGCGCGCGC 3043

QY      41  ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      3044  ATCTCGAGCAGCGCGCGCAGCGCGCTCGCGCGCGCTCGACGGAGGCGAGCGAGCGC 3103

QY      61  LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      3104  GTGCGGCGAGCGCGAGGAGGCGCGCGCGCTCTGACGGAGGGCGAGGAG 3151

RESULT 7
US-10-172-118-1743
; Sequence 1743, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1743
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_018104
; DATABASE ENTRY DATE: 2001-06-18
; US-10-172-118-1743

```

```

Alignment Scores:
Pred. No.:      1.64e-33      Length:      3614
Score:          296.00      Matches:      54
Percent Similarity: 86.84%      Conservative: 12
Best local Similarity: 71.05%      Mismatches:  10
Query Match:      75.70%      Indels:      0
DB:              13          Gaps:        0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-172-118-1743 (1-3614)

QY      1  SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      2924  TGGCAGTCCACACCGGTGGTGAACGGCAGACGGCGCAGGTTCCGGACAGTGGAGATGCAG 2983
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY      21  TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      2984  TTCGGCGGCTGGCGCTGCACGTGCGTACGGCATGCCCTGGACGAGGAGAGCGCGC 3043
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY      41  ValLeuGluLeuAlaArgGlnArgAlaValargGlnAlaTyrAlaArgGluGlnArg 60
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      3044  ATCTCTGGACGAGCGCGCAGCGCGCTCTCCCGCGGCTGGCGCGCGGACGACGAGCGC 3103
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

Qy 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 3104 GTGGCGACGCGAGGAGCGCGCGCTCTGGACGAGGCGGAG 3151

RESULT 8

US-10-038-854-39

; Sequence 39, Application US/10038854

; Publication No. US20040022781A1

; GENERAL INFORMATION:

; APPLICANT: Spytek, Kimberly A

; APPLICANT: Li, Li

; APPLICANT: Wolenc, Adam R

; APPLICANT: Vernet, Corine

; APPLICANT: Eisen, Andrew J

; APPLICANT: Liu, Xiaohong

; APPLICANT: Malyankar, Uriel M

; APPLICANT: Shimkets, Richard A

; APPLICANT: Tchernev, Velizar

; APPLICANT: Spaderna, Steven K

; APPLICANT: Gorman, Linda

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Patturajan, Meera

; APPLICANT: Gusev, Vladimir Y

; APPLICANT: Gangolli, Esha A

; APPLICANT: Guo, Xiaojia S

; APPLICANT: Shenoy, Suresh G

; APPLICANT: Rastelli, Luca

; APPLICANT: Casman, Stacie J

; APPLICANT: Boldog, Ferenc

; APPLICANT: Burgess, Catherine E

; APPLICANT: Edinger, Shlomit R

; APPLICANT: Ellerman, Karen

; APPLICANT: Gunther, Erik

; APPLICANT: Smithson, Glenda

; APPLICANT: Millet, Isabelle

; APPLICANT: MacDougall, John R

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-230

; CURRENT APPLICATION NUMBER: US/10/038,854

; PRIOR FILING DATE: 2003-01-22

; PRIOR APPLICATION NUMBER: 60/258,928

; PRIOR FILING DATE: 2000-12-29

; PRIOR APPLICATION NUMBER: 60/259,415

; PRIOR FILING DATE: 2001-01-02

; PRIOR APPLICATION NUMBER: 60/259,785

; PRIOR FILING DATE: 2001-01-04

; PRIOR APPLICATION NUMBER: 60/269,814

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/279,832

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,833

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,863

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/283,889

; PRIOR FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: 60/284,447

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/286,683

; PRIOR FILING DATE: 2001-04-25

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 411

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 39

; LENGTH: 8473

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-038-854-39

Alignment Scores:

Pred. No.: 4,89e-33

Score: 296.00

Matches: 86.84%

Percent Similarity:

Length: 8473

Matches: 54

Conservative: 12

Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-038-854-39 (1-8473)

Qy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 7782 TCGCAGTCCACACGCTGGTGAACGCGAGGACGGCGAGTTCGGCGACGTGGAGATGCAG 7841

Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 7842 TTCGGCGCGCTGGCGCTACGCTGCGTACGCGACCTCGGACGAGGAGGCGCGC 7901

Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db 7902 ATCCTCGAGGAGGCGCGCGCGCTCGCCCGGCGCTCGCGCGGCGCGGCGGAGCGCGC 7961

Qy 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 7962 GTGGCGACGCGGAGGAGGCGCGCGCTCTGGACGAGGCGGAGGAG 8009

RESULT 9

US-10-038-854-41

; Sequence 41, Application US/10038854

; Publication No. US20040022781A1

; GENERAL INFORMATION:

; APPLICANT: Spytek, Kimberly A

; APPLICANT: Li, Li

; APPLICANT: Wolenc, Adam R

; APPLICANT: Vernet, Corine

; APPLICANT: Eisen, Andrew J

; APPLICANT: Liu, Xiaohong

; APPLICANT: Malyankar, Uriel M

; APPLICANT: Shimkets, Richard A

; APPLICANT: Tchernev, Velizar

; APPLICANT: Spaderna, Steven K

; APPLICANT: Gorman, Linda

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Patturajan, Meera

; APPLICANT: Gusev, Vladimir Y

; APPLICANT: Gangolli, Esha A

; APPLICANT: Guo, Xiaojia S

; APPLICANT: Shenoy, Suresh G

; APPLICANT: Rastelli, Luca

; APPLICANT: Casman, Stacie J

; APPLICANT: Boldog, Ferenc

; APPLICANT: Burgess, Catherine E

; APPLICANT: Edinger, Shlomit R

; APPLICANT: Ellerman, Karen

; APPLICANT: Gunther, Erik

; APPLICANT: Smithson, Glenda

; APPLICANT: Millet, Isabelle

; APPLICANT: MacDougall, John R

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-230

; CURRENT APPLICATION NUMBER: US/10/038,854

; PRIOR FILING DATE: 2003-01-22

; PRIOR APPLICATION NUMBER: 60/258,928

; PRIOR FILING DATE: 2000-12-29

; PRIOR APPLICATION NUMBER: 60/259,415

; PRIOR FILING DATE: 2001-01-02

; PRIOR APPLICATION NUMBER: 60/259,785

; PRIOR FILING DATE: 2001-01-04

; PRIOR APPLICATION NUMBER: 60/269,814

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/279,832

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,833

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,863

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/283,889

FILE REFERENCE: 21402-230
CURRENT APPLICATION NUMBER: US/10/038,854
CURRENT FILING DATE: 2003-01-22
PRIOR APPLICATION NUMBER: 60/258,928
PRIOR FILING DATE: 2000-12-29
PRIOR APPLICATION NUMBER: 60/259,415
PRIOR FILING DATE: 2001-01-02
PRIOR APPLICATION NUMBER: 60/259,785
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: 60/269,814
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/279,832
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,833
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,863
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/283,889
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,447
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/286,683
PRIOR FILING DATE: 2001-04-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 411
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 37
LENGTH: 8645
TYPE: DNA
ORGANISM: Homo sapiens
US-10-038-854-37

Alignment Scores:
Pred. No.: 5 02e-33 Length: 8645
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservative: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-038-854-37 (1-8645)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 7954 TCGCAGTCCACCACCGGTGGTGAACGCGAGCGCGAGGTTTCGCGGACGTGGAGATGCGAG 8013

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluLysAlaArg 40
Db 8014 TTCGGCGCGCTGGCGCTGCACGTGGCTACGGCATGACCTGGACGAGGAGGCGCGC 8073

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 8074 ATCCTGGAGCAGCGCGCGCGCGCTCGCCCGCGCTCGCCCGCGCGAGCAGCGC 8133

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8134 GTGCGCGAGCGCGAGGAGGCGCGCGCTCTGGACGAGGCGGAGGAG 8181

RESULT 11
US-10-038-854-35
Sequence 35, Application US/10038854
Publication No. US20040022781A1
GENERAL INFORMATION:
APPLICANT: Spytex, Kimberly A
APPLICANT: Li, Li
APPLICANT: Wolenc, Adam R
APPLICANT: Vernet, Corine
APPLICANT: Eisen, Andrew J
APPLICANT: Liu, Xiaohong
APPLICANT: Malyankar, Uriel M
APPLICANT: Shimkets, Richard A
APPLICANT: Tchernev, Velizar
APPLICANT: Spaderna, Steven K

PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,447
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/286,683
PRIOR FILING DATE: 2001-04-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 411
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 8487
TYPE: DNA
ORGANISM: Homo sapiens
US-10-038-854-41

Alignment Scores:
Pred. No.: 4 9e-33 Length: 8487
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservative: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-038-854-41 (1-8487)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 7778 TCGCAGTCCACCACCGGTGGTGAACGCGAGCGCGAGGTTTCGCGGACGTGGAGATGCGAG 7837

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluLysAlaArg 40
Db 7838 TTCGGCGCGCTGGCGCTGCACGTGGCTACGGCATGACCTGGACGAGGAGGCGCGC 7897

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 7898 ATCCTGGAGCAGCGCGCGCGCGCTCGCCCGCGCTCGCCCGCGCGAGCAGCGC 7957

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 7958 GTGCGCGAGCGCGAGGAGGCGCGCGCTCTGGACGAGGCGGAGGAG 8005

RESULT 10
US-10-038-854-37
Sequence 37, Application US/10038854
Publication No. US20040022781A1
GENERAL INFORMATION:
APPLICANT: Spytex, Kimberly A
APPLICANT: Li, Li
APPLICANT: Wolenc, Adam R
APPLICANT: Vernet, Corine
APPLICANT: Eisen, Andrew J
APPLICANT: Liu, Xiaohong
APPLICANT: Malyankar, Uriel M
APPLICANT: Shimkets, Richard A
APPLICANT: Tchernev, Velizar
APPLICANT: Spaderna, Steven K
APPLICANT: Gorman, Linda
APPLICANT: Kekuda, Ramesh
APPLICANT: Patturajan, Meera
APPLICANT: Gusev, Vladimir Y
APPLICANT: Gangoli, Esha A
APPLICANT: Guo, Xiaojia S
APPLICANT: Shenoy, Suresh G
APPLICANT: Rastelli, Luca
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc
APPLICANT: Burgess, Catherine E
APPLICANT: Edinger, Shlomit R
APPLICANT: Ellerman, Karen
APPLICANT: Gunther, Erik
APPLICANT: Smithson, Glennda
APPLICANT: Millet, Isabelle
APPLICANT: MacDougall, John R
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

```

; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimiro Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-038-854-35

Alignment Scores:
Pred. No.: 5,04e-33 Length: 8675
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservative: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-038-854-35 (1-8675)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 7966 TCCAGTCCACACGGTGGTGAACGCGAGCGGAGTTCGCGACGTGGAGATGCAG 8025
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGlySalaArg 40
Db 8026 TTCGGCGCGTGGCGTGCACGCGTACCGCATCACCTGCGACGAGGAGGCGCG 8085
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 8086 ATCCTGAGCAGCGCGCGCGCTGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCG 8145
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
```

```

Db 8146 GTGCGGACGCGCGAGGAGGGCGCGCGCTCTGGACGAGCGCGAGAAG 8193
RESULT 12
US-09-808-602-79
; Sequence 79, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Heriman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 79
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
; US-09-808-602-79

Alignment Scores:
Pred. No.: 4.24e-26 Length: 8409
Score: 249.50 Matches: 48
Percent Similarity: 75.95% Conservative: 12
Best Local Similarity: 60.76% Mismatches: 16
Query Match: 63.81% Indels: 3
DB: 9 Gaps: 1

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-808-602-79 (1-8409)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 8038 TCCAGCGCAACCTCTCTTATCAACGGAAGACTCGACGGTTTCAAAATCGAGTTTCAG 8097
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu 37
Db 8098 TATTCACCTGCTGTATCAACATCCGCTACGGGCTCACCGCGACACGCTGGATGAGGAG 8157
QY 38 LysAlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGlu 57
Db 8158 AAGGCGACGAGTGTAGACAGGCTCGGCGAGCGAGCCCTGGGGTCGCGCTGGGCCAAAGAG 8217
QY 58 GlnGlnArgLeuArgGluGluGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8218 CACGAGAGGCGAGGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 8274
RESULT 13
US-09-800-198-67
; Sequence 67, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
```

```

; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 67
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-800-198-67

Alignment Scores:
Pred. No.: 8409
Score: 249.50
Percent Similarity: 75.95%
Best Local Similarity: 60.76%
Query Match: 63.81%
Indels: 3
Gaps: 1

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-800-198-67 (1-8409)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 1717 TCCAGCCACCGCTGCTGCTCAACCGCAGGACTCGAAGTTTCAGAACTGAGTTCCAG 1658
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu 37
Db 1657 TACTCCAGCGTCTGCTCAGCATCGCTATGCGCTCACCCCGCACACCTGGACGAAGAG 1598
QY 38 LysAlaArgValLeuGluLeuAlaArgGlnAlaValArgGlnAlaAlaTrpAlaArgGlu 57
Db 1597 AAGCGCCGCTCTGACCGAGAGAGAGGCGGCGGAGCGGCGCTGTGGACTGAGGCGAGAAG 1538
QY 58 GlnGlnArgLeuArgGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 76
Db 1537 CAGCAGAAAGCCAGGACGGAGAGAGAGGCGGCGGAGCGGCGCTGTGGACTGAGGCGAGAAG 1481

RESULT 15
US-09-808-602-76
; Sequence 76, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Khumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 76
; LENGTH: 6560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-808-602-76

Alignment Scores:
Pred. No.: 8,64e-26
Score: 246.50
Percent Similarity: 75.95%
Best Local Similarity: 62.03%
Query Match: 63.04%
Indels: 3
Gaps: 1

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-808-602-76 (1-6560)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 4844 TCCAGCCACCGCTGCTGCTCAACCGCAGGACTCGAAGTTTCAGAACTGAGTTCCAG 4903
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu 37
Db 4904 TACTCCAGCGTCTGCTCAGCATCGCTATGCGCTCACCCCGCACACCTGGACGAAGAG 4963
QY 38 LysAlaArgValLeuGluLeuAlaArgGlnAlaValArgGlnAlaAlaTrpAlaArgGlu 57
Db 4964 AAGGCGCGCTCTGACCGAGAGAGGCGGCGGAGCGGCGCTGGCGACGCGCTGGGCGCAAGAG 5023
QY 58 GlnGlnArgLeuArgGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 76
Db 5024 CAGCAGAAAGCCAGGACGGAGAGAGGCGGCGGAGCGGCGCTGTGGACTGAGGCGAGAAG 5080

Search completed: August 14, 2004, 19:23:12
```

```

; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 67
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-800-198-67

Alignment Scores:
Pred. No.: 8409
Score: 249.50
Percent Similarity: 75.95%
Best Local Similarity: 60.76%
Query Match: 63.81%
Indels: 3
Gaps: 1

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-800-198-67 (1-8409)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 8038 TCCAGCCACCGCTGCTGCTCAACCGAAGGACTCGAAGTTTCAGAACTGAGTTCCAG 8097
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu 37
Db 8098 TATTCCAGCGTCTGCTCAGCATCGCTACGGGCTCACCGCGACACGCTGGATGAGGAG 8157
QY 38 LysAlaArgValLeuGluLeuAlaArgGlnAlaValArgGlnAlaAlaTrpAlaArgGlu 57
Db 8158 AAGGACGAGTCTGACGAGGCTCGGACGAGGCGGCTGGGCTGGGCGGCGGCGGCGGCGG 8217
QY 58 GlnGlnArgLeuArgGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 76
Db 8218 CAGCAGAAAGCCAGGACGGAGAGAGGCGGCGGAGCGGCGCTGTGGACTGAGGCGAGAAG 8274

RESULT 14
US-10-198-846-13976/c
; Sequence 13976, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13976
; LENGTH: 3217
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-13976

Alignment Scores:
Pred. No.: 3,46e-26
Score: 246.50
Percent Similarity: 75.95%
Best Local Similarity: 62.03%
Query Match: 63.04%
Indels: 3
Gaps: 1

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-198-846-13976 (1-3217)
```


This Page Blank (uspio)

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-223-134-1 (1-14113)


```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2453
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2453

Alignment Scores:
Pred. No.: 7.67 Length: 1008
Score: 68.00 Matches: 24
Percent Similarity: 50.72% Conservative: 11
Best Local Similarity: 34.78% Mismatches: 15
Query Match: 17.39% Indels: 19
DB: 4 Gaps: 3

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-252-991A-2453 (1-1008)
Qy 15 ThrAspIleGlnLeuGlnTyrGlyAla-----LeuCysLeuAsnThrArgTyrGlyThr 32
Db 809 ACCCTGCTCAAGCTCGCCTCAGGCAGTCCCGAGGCGCTGCATTACCC----- 856
Qy 33 ThrLeuAspGluGluLysAlaAArg-----ValLeuGlu 43
Db 857 -----GAGGAACTGGCGCGGAATGCGCCCTGGCTGGGAAGTGGCCCTGAG 907
Qy 44 LeuAlaArgGln-ArgAlaValArgGlnAlaTrrPalaArgGluGlnArgLeuArgGl 63
Db 908 CTGCGCCAGGAGAGAGGGCCCGCGCCGATCTCGCGCAGGACTACGAAAGTCACTGCGGAA 967
Qy 63 uGlyGluGluGlyLeuArgAlaTrrP 71
Db 968 ATGCTTCGCAAGCTCCGGAATGG 992

RESULT 11
US-09-252-991A-12249
; Sequence 12249, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12249
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12249

Alignment Scores:
Pred. No.: 11.6 Length: 1392
Score: 68.00 Matches: 20
Percent Similarity: 44.83% Conservative: 6
Best Local Similarity: 34.48% Mismatches: 32
Query Match: 17.39% Indels: 0
DB: 4 Gaps: 0

```

```
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 2535
/ LENGTH: 2205
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2535

Alignment Scores:
Pred. No.: 20.8      Length: 2205
Score: 68.00        Matches: 24
Percent Similarity: 50.72%      Conservative: 11
Best Local Similarity: 34.78%    Mismatches: 15
Query Match: 17.39%             Indels: 19
DB: 4                      Gaps: 3

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-252-991A-2535 (1-2205)

QY 15 ThrAspIleGlnLeuGlnTyrGlyAla-----LeuCysLeuAsnThrArgTyrGlyThr 32
Db 1276 ACCCTGCTCAAGCTCGCTACGGCAGTCCCGAGGCTCGGTATTACCC-----1323
QY 33 ThrLeuAspGluGluLysAlaArg-----ValLeuGlu 43
Db 1324 -----GAGGAAGTGGCGCGGAAATGGCCCTGGTGGGAAATGGCCCTGGAG 1374
QY 44 LeuAlaArgGln-ArgAlaValArgGlnAlaTrpAlaArgGluGlnArgLeuArgGlu 63
Db 1375 CTGGCCAGGAGAGAGGCGCGCGGATCTGGCGCAGGACTACGAAGTCACTCGGAA 1434
QY 63 uGlyGluGluGlyLeuArgAlaTrp 71
Db 1435 ATGCTTCGCAAGCGTCCGGAATGG 1459

RESULT 14
US-09-252-991A-2384/c
/ Sequence 2384, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 2384
/ LENGTH: 2388
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2384

Alignment Scores:
Pred. No.: 23      Length: 2388
Score: 68.00        Matches: 24
Percent Similarity: 50.72%      Conservative: 11
Best Local Similarity: 34.78%    Mismatches: 15
Query Match: 17.39%             Indels: 19
DB: 4                      Gaps: 3

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-252-991A-2384 (1-2388)

QY 15 ThrAspIleGlnLeuGlnTyrGlyAla-----LeuCysLeuAsnThrArgTyrGlyThr 32
Db 1095 ACCCTGCTCAAGCTCGCTACGGCAGTCCCGAGGCTCGGTATTACCC-----1048
QY 33 ThrLeuAspGluGluLysAlaArg-----ValLeuGlu 43
Db 1047 -----GAGGAAGTGGCGCGGAAATGGCCCTGGTGGGAAATGGCCCTGGAG 997
```

```
QY 44 LeuAlaArgGln-ArgAlaValArgGlnAlaTrpAlaArgGluGlnArgLeuArgGlu 63
Db 996 CTGGCCAGGAGAGAGGCGCGCGGATCTGGCGCAGGACTACGAAGTCACTCGGAA 937
QY 63 uGlyGluGluGlyLeuArgAlaTrp 71
Db 936 ATGCTTCGCAAGCGTCCGGAATGG 912
```

```
RESULT 15
US-09-621-976-595
/ Sequence 595, Application US/09621976
/ Patent No. 6639063
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Jobert, S.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: ESTs and Encoded Human Proteins.
/ FILE REFERENCE: GENSET.054PR2
/ CURRENT APPLICATION NUMBER: US/09/621,976
/ CURRENT FILING DATE: 2000-07-21
/ NUMBER OF SEQ ID NOS: 19335
/ SOFTWARE: Patent.pm
/ SEQ ID NO 595
/ LENGTH: 374
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 199..372
US-09-621-976-595
```

```
Alignment Scores:
Pred. No.: 2.93      Length: 374
Score: 67.00         Matches: 15
Percent Similarity: 68.00%      Conservative: 2
Best Local Similarity: 60.00%    Mismatches: 6
Query Match: 17.14%             Indels: 2
DB: 4                      Gaps: 1
```

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-621-976-595 (1-374)

```
QY 46 ArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArgLeuArgGluGlyGlu 65
Db 19 CGGAGAGGCGCACTCGCGATGATGTCGTCGG-----CAGGCTCTTCGGGAAGGGAA 72
QY 66 GluGlyLeuArgAla 70
Db 73 GAGGAGAAAGGGCC 87
```

Search completed: August 14, 2004, 21:40:39
Job time : 34.4845 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 891.313 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_2650_2725
Perfect score: 391
Sequence: 1 SQINTVLNGRRTRYTDIQLQ.....PQRLREGEGGLRAWTEGEK 76

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112216_29287/app.query.fasta_1.3519
-DB=EST -QPMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029020@cgn 1 13135 @runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_esthum:*
4: em_esthum:*
5: em_estov:*
6: em_estpl:*
7: em_estio:*
8: em_htc:*
9: gb_esti:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pig:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 391 | 100.0 | 6246 | 29 | AY413475 Homo sapi |
| 2 | 385 | 98.5 | 499 | 10 | BF077525 227453 MA |
| 3 | 385 | 98.5 | 685 | 13 | BU742156 UI-E-EJ0- |
| 4 | 384 | 98.2 | 5970 | 29 | AY413476 Pan trogl |
| 5 | 371 | 94.9 | 881 | 13 | BU384298 603858267 |
| 6 | 371 | 94.9 | 976 | 13 | BU384298 603858267 |
| 7 | 368 | 94.1 | 621 | 14 | CF182790 UI-M-EY0- |
| 8 | 365 | 93.4 | 913 | 13 | BU387427 603857864 |
| 9 | 364 | 93.1 | 693 | 13 | BQ210132 UI-R-EP0- |
| 10 | 358 | 91.6 | 514 | 9 | AI176158 EST19739 |
| 11 | 353 | 90.3 | 424 | 9 | AV589580 AV589580 |
| 12 | 344 | 88.0 | 748 | 9 | AU169909 AV589580 |
| 13 | 335 | 85.7 | 507 | 10 | AW580211 RC4-HT041 |
| 14 | 313 | 80.1 | 713 | 9 | AL045769 DKFZP434F |
| 15 | 303 | 77.5 | 608 | 10 | BB161536 BB161536 |
| 16 | 303 | 77.5 | 949 | 29 | CNS03WBO Tetraodon |
| 17 | 303 | 77.5 | 3038 | 11 | AK037897 Mus muscu |
| 18 | 299 | 76.5 | 769 | 13 | BU454148 603771046 |
| 19 | 299 | 76.5 | 863 | 13 | BU382750 603859889 |
| 20 | 299 | 76.5 | 946 | 13 | BU132264 603120954 |
| 21 | 296 | 75.7 | 311 | 9 | AI184444 qd40d07.x |
| 22 | 296 | 75.7 | 359 | 9 | AI222954 qm30d03.x |
| 23 | 296 | 75.7 | 378 | 9 | AI200670 qf94c12.x |
| 24 | 296 | 75.7 | 380 | 9 | AI887009 w195d07.x |
| 25 | 296 | 75.7 | 382 | 9 | AI199810 qf95c06.x |
| 26 | 296 | 75.7 | 382 | 9 | AI274285 ql61c09.x |
| 27 | 296 | 75.7 | 382 | 12 | BM171993 UI-E-EJ0- |
| 28 | 296 | 75.7 | 390 | 9 | AI144798 UI-R-BT0- |
| 29 | 296 | 75.7 | 390 | 9 | AW044155 w170e05.x |
| 30 | 296 | 75.7 | 394 | 9 | AI910459 w131c06.x |
| 31 | 296 | 75.7 | 395 | 9 | AI310370 qo66h01.x |
| 32 | 296 | 75.7 | 401 | 9 | AI250756 q136b04.x |
| 33 | 296 | 75.7 | 401 | 9 | AI367099 qk66e03.x |
| 34 | 296 | 75.7 | 405 | 10 | AW137000 UI-H-B11- |
| 35 | 296 | 75.7 | 406 | 9 | AI911042 wd22g06.x |
| 36 | 296 | 75.7 | 413 | 10 | AW207016 UI-H-B11- |
| 37 | 296 | 75.7 | 417 | 9 | AI262318 qk34c03.x |
| 38 | 296 | 75.7 | 421 | 9 | AI391741 qy94c12.x |
| 39 | 296 | 75.7 | 430 | 9 | AI338238 qt41h06.x |
| 40 | 296 | 75.7 | 440 | 10 | AW137494 UI-H-B11- |
| 41 | 296 | 75.7 | 447 | 9 | AI827404 wk94b10.x |
| 42 | 296 | 75.7 | 450 | 9 | AI288964 ql81d06.x |
| 43 | 296 | 75.7 | 453 | 9 | AI278290 ql178d04.x |
| 44 | 296 | 75.7 | 454 | 9 | AI079962 oz34g12.x |
| 45 | 296 | 75.7 | 458 | 9 | AI017896 ou35g05.x |

ALIGNMENTS

RESULT 1
AY413475
LOCUS
DEFINITION
Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
6246 bp DNA linear GSS 12-DEC-2003
genomic survey sequence.
ACCESSION
AY413475
VERSION
AY413475.1
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 6246)

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Dr. M. Bento Soares, University of Iowa
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source
1..685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-aim-m-24-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"
/dev_stages="fetal and adult"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldi, Lennan and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGATCAAGCA; lens, CGATTAGCGA; eye anterior segment, AATGCCGACAT; optic nerve, CCATTAAAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). TAG_TISSUB=human fetal eyes
TAG_LIB=UI-E-EJ0
TAG_SEQ=AGATCAAGA"

ORIGIN

Alignment Scores:
Pred. No.: 1.79e-33 Length: 685
Score: 385.00 Matches: 75
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 1
Query Match: 98.47% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x BU742156 (1-685)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 640 TCCAGATCAACACAGTACTTATGCCAGACTAGACNTACAGCATCCAGCTCCAG 581
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 580 TACGGGGCACTGTGCTTGAACACACAGCTACGGACACAGTTGGATGAGGAGGACCG 521
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 520 GTCTCTGAGCTGGCCCGGCGAGAGCGGTGCGCAAGCGTGGCGCGCGAGCAGAGAGA 461

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 460 CTGCGGGAAGGGGAGGAGGCGCTGCGGGCTGGACAGAGGGGAGAG 413
RESULT 4
AY413476
LOCUS
DEFINITION
Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY413476
VERSION
AY413476.1 GI:39769438
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 5970)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 5970)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..5970
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..5970
/locus_tag="HCM4903"
ORIGIN
Alignment Scores:
Pred. No.: 5.3e-32 Length: 5970
Score: 384.00 Matches: 75
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 1
Query Match: 98.21% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_2650_2725 (1-76) x AY413476 (1-5970)
QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 5608 TCCAGATCAACACAGTACTTATGCCAGACTAGACGCTACAGCATCCAGCTCCAG 5667
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 5668 TACGGGGCACTGTGCTTGAACACACAGCTACGGACACAGTTGGATGAGGAGGACCG 5727
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 5728 GTCTCTGAGCTGGCCCGGCGAGAGCGGTGCGCAAGCATGGCCCGGCGAGCAGAGA 5787
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 5788 CTGCGGGAANNNGAGGAGGCGCTGCGGGCTGGACAGAGGGGAGAGAG 5835
RESULT 5
BU384298

LOCUS BU384298 881 bp mRNA linear EST 28-NOV-2002
DEFINITION 603858267F1 CSEQCHN75 Gallus gallus cDNA clone CHEST865f18 5', mRNA
sequence.
ACCESSION BU384298
VERSION BU384298.1 GI:25892299
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 881)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1..881
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hixex"
/db_xref="taxon:9031"
/clone="CHEST865f18"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN75"
/note="Organ: trunk; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bernaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.: 9,93e-32 Length: 881
Score: 371.00 Matches: 71
Percent Similarity: 97.37% Conservativeness: 3
Best Local Similarity: 93.42% Mismatches: 2
Query Match: 94.88% Indels: 0
DB: 13 Gaps: 0
US-10-029-020-14_COPY_2650_2725 (1-76) x BU384298 (1-881)

Qy 1 SerGlnIleAenThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 47 TCCAGATCAACACAGTCTGGTGAAGGAGTACAGCTTACAGGACATCCAGCTGAG 106
Qy 21 TyrGlyAlaLeuCyLeuAsnThrArgGlyThrLeuAspGluGluLeuAlaArg 40
Db 107 TATGTGCGCTGTGTCTGAAACATCTGCTACGGGACCACTTGGACGAGGAGGCCCGGA 166
Qy 41 ValLeuGluLeuAlaAraGlnAraAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 167 GTCCTGGAGCTGGCCCGGCGGCGCGCTGGCCCAAGCTTGGCCCGGGAACAGCAGAGA 226

Qy 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 227 TTGAGGGATGGGAGGAGGATTCGCTCGTGGACAGAGGGAGAG 274
RESULT 6
LOCUS BU365499 976 bp mRNA linear EST 28-NOV-2002
DEFINITION 603788050F1 CSEQCHN72 Gallus gallus cDNA clone CHEST747c13 5', mRNA
sequence.
ACCESSION BU365499
VERSION BU365499.1 GI:25873500
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 976)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1..976
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST747c13"
/sex="Female"
/tissue_type="cerebrum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN72"
/note="Organ: brain; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bernaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.: 1.15e-31 Length: 976
Score: 371.00 Matches: 71
Percent Similarity: 97.37% Conservativeness: 3
Best Local Similarity: 93.42% Mismatches: 2
Query Match: 94.88% Indels: 0
DB: 13 Gaps: 0
US-10-029-020-14_COPY_2650_2725 (1-76) x BU365499 (1-976)

Qy 1 SerGlnIleAenThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 47 TCCAGATCAACACAGTCTGGTGAAGGAGTACAGCTTACAGGACATCCAGCTGAG 106
Qy 21 TyrGlyAlaLeuCyLeuAsnThrArgGlyThrLeuAspGluGluLeuAlaArg 40
Db 107 TATGTGCGCTGTGTCTGAAACATCTGCTACGGGACCACTTGGACGAGGAGGCCCGGA 166
Qy 41 ValLeuGluLeuAlaAraGlnAraAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 167 GTCCTGGAGCTGGCCCGGCGGCGCGCTGGCCCAAGCTTGGCCCGGGAACAGCAGAGA 226

```

Db      173  TCCAGATCAACACAGCTGCTGGTGAAGGACTAGACGTTACACGAGCATCCAGCTGCAG 232
QY      21   TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db      233  TACGGTGCCTGTGTCTGAACACTCGCTACGGAGCACCTTTGACGAGGAGGAGGCCGCA 232
QY      41   ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrrpAlaArgGluGlnGlnArg 60
Db      293  GTCTTGAGCTGGCCCGCAGCGCGCGTGGCCCAAGCTTGGCCCGGGAACAGCAGAGA 352
QY      61   LeuArgGluGlyGluGluGlyLeuArgAlaTrrpThrGluGlyGluLys 76
Db      353  TTGAGGATGGGAGGAGGAGGATTCGCTCGTGACACAGAGGGGAGAG 400

RESULT 7
LOCUS   CF182790
DEFINITION   UI-M-EYO-bwz-p-16-0-UI.r1 NIH BMAP EYO Mus musculus cDNA clone
IMAGE: 5705343 5', mRNA sequence.
ACCESSION   CF182790
VERSION     CF182790.1 GI:33314674
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus

REFERENCE
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL     NIH-MGC http://mgi.nci.nih.gov/
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: Dr. James Lin, University of Iowa
            cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)

Seq primer: pyX-5.
Location/Qualifiers
1. .621
   /organism="Mus musculus"
   /mol_type="mRNA"
   /strain="C57BL/6"
   /db_xref="taxon:10090"
   /clone="IMAGE: 5705343"
   /tissue_type="whole brain"
   /dev_stage="embryo 15.5 dpc"
   /lab_host="DH10B (T1 phage resistant)"
   /clone_lib="NIH_BMAP_EYO"
   /note="Organ: brain; Vector: pyX-Asc; Site 1: Ecor I;
         Site 2: Not I; The library was constructed according to
         Ronaldo, Lennon and Soares, Genome Research, 6:791-806,
         1996. Denatured mRNA was size fractionated on a 1% agarose
         gel. First strand cDNA synthesis was primed with an
         oligo-dT primer containing a Not I site. Double stranded
         cDNA was size selected according to mRNA size fraction,
         ligated with Ecor I adaptor, digested with Not I, and then
         cloned directionally into pyX-Asc vector. The library tag
         sequence located between the Not I site and the polyA
         tail, is GTGGGTGAA. This library was created for the
         University of Iowa Mouse Brain Molecular Anatomy Project
         (BMAP): 'Gene Discovery in the Developing Mouse Nervous
         System', supported by National Institutes of Mental Health
         (NIMH), Hemin Chin, Ph.D., program coordinator."

FEATURES
source
1. .621
   /organism="Mus musculus"
   /mol_type="mRNA"
   /strain="C57BL/6"
   /db_xref="taxon:10090"
   /clone="IMAGE: 5705343"
   /tissue_type="whole brain"
   /dev_stage="embryo 15.5 dpc"
   /lab_host="DH10B (T1 phage resistant)"
   /clone_lib="NIH_BMAP_EYO"
   /note="Organ: brain; Vector: pyX-Asc; Site 1: Ecor I;
         Site 2: Not I; The library was constructed according to
         Ronaldo, Lennon and Soares, Genome Research, 6:791-806,
         1996. Denatured mRNA was size fractionated on a 1% agarose
         gel. First strand cDNA synthesis was primed with an
         oligo-dT primer containing a Not I site. Double stranded
         cDNA was size selected according to mRNA size fraction,
         ligated with Ecor I adaptor, digested with Not I, and then
         cloned directionally into pyX-Asc vector. The library tag
         sequence located between the Not I site and the polyA
         tail, is GTGGGTGAA. This library was created for the
         University of Iowa Mouse Brain Molecular Anatomy Project
         (BMAP): 'Gene Discovery in the Developing Mouse Nervous
         System', supported by National Institutes of Mental Health
         (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 1.31e-31 Length: 621
Score: 368.00 Matches: 71

```

```

Percent Similarity: 97.37%
Best Local Similarity: 93.42%
Query Match: 94.12%
DB: 14
Conservative: 3
Mismatch: 2
Indels: 0
Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x CF182790 (1-621)

QY      1   SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db      377  TCCAGATCAACACCGTGTCTAGTGGCAGGACTAGACGCTACACTGACATCCAGCTGCAG 436
QY      21   TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db      437  TACAGGGCGCTGTGCTGAACACCCGCTACGGGCAACACGTCGATGAGGAAAGGTGCG 496
QY      41   ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrrpAlaArgGluGlnArg 60
Db      497  GTGCTGGAGCTGGCCAGGAGAGAGCTGTGCCAGGCTTGGCCCGGAGAGCAGAGA 556
QY      61   LeuArgGluGlyGluGluGlyLeuArgAlaTrrpThrGluGlyGluLys 76
Db      557  CTGCGGGAAGGGGAAGAGGGCTGCGGGCTGACAGATGGGGAAG 604

RESULT 8
LOCUS   BU387427
DEFINITION   603857864F1 CSEQHN75 Gallus gallus cDNA clone ChEST864p16 5', mRNA
IMAGE: 603857864F1 CSEQHN75 Gallus gallus cDNA clone ChEST864p16 5', mRNA
ACCESSION   BU387427
VERSION     BU387427.1 GI:25895415
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
ORGANISM    Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 913)
AUTHORS     Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
            A Comprehensive Collection of Chicken cDNAs
            Curr. Biol. 12 (22), 1965-1969 (2002)
            22335534
            12445392
            Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612088930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
            Location/Qualifiers
1. .913
   /organism="Gallus gallus"
   /mol_type="mRNA"
   /strain="White Leghorn, HiseX"
   /db_xref="taxon:9031"
   /clone="ChEST864p16"
   /dev_stage="36"
   /lab_host="DH10B"
   /clone_lib="CSEQHN75"
   /note="Organ: trunks; Vector: pBluescript II KS(+);
         Site 1: EcoRI; Site 2: NotI; This normalized library was
         constructed from 1 million independent clones. cDNA
         synthesis was initiated using an oligo(dT) primer, using
         methylated C in the first strand synthesis reaction.
         Following this first strand reaction, double-stranded cDNA
         was blunted, ligated to NotI adapters, digested with
         EcoRI, size-selected, and cloned into the NotI and EcoRI
         compatible sites of a custom modified MCS of the
         pBluescript (KS+) vector. The library was normalized in 2
         rounds using conditions adapted from Soares et al., PNAS
         (1994) 91: 9228-9232 and Ronaldo et al., Genome Research 6

```

(1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores: 5e-31 Length: 913
Pred. No.: 365.00 Matches: 70
Score: 96.05% Conservative: 3
Percent Similarity: 92.11% Mismatches: 3
Best Local Similarity: 93.35% Indels: 0
Query Match: 13 Gaps: 0
DB:

US-10-029-020-14_COPY_2650_2725 (1-76) x BU387427 (1-913)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 111 TCCAGATCAACACAGTGTGGTGAAGACTAGACGTTACACGACATCCAGCTGCAG 170
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluLysAlaArg 40
Db 171 TATGGTGGCTGTGTCTGAACACTCGCTACGGGACCACITTTGACGAGGAGAGACCA 230
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 231 GTCTTGAGCTGGCCCGGACGCGCGGTGGCCCAAGCTTGGCCCGGACACAGCAGA 290
QY 61 LeuArgGluGluGluGluGluArgAlaTrpThrGluGlyGluLys 76
Db 291 TTGAGGATGGGAGGAGGGGATTCGCTCGTGACAGAGGGGAGAAG 338

RESULT 9

BQ210132/c 693 bp mRNA linear EST 02-MAY-2002
LOCUS UI-R-EP0-cod-1-03-0-UI.s1 UI-R-EP0 Rattus norvegicus cDNA clone
DEFINITION UI-R-EP0-cod-1-03-0-UI 3', mRNA sequence.

ACCESSION BQ210132 GI:20426597
VERSION BQ210132.1
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 693)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548

COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9585
Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized osteoblast library cDNA library preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLVA=No.

FEATURES

source 1..693
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"

/clone="UI-R-EP0-cod-1-03-0-UI"
/dev stage="ADULT"
/lab host="DH10B (Life Technologies)"
/clone lib="UI-R-EP0"
/note="Vector: pMT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; UI-R-EP0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pMT3D-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: distal colon, GAAAGTGCTCC; osteoblast, AAGATATCAA; cell line R3327 5A, GGACTAGATC; cell line R3327 5P, CACGTGAGAT; duodenum, TGTGGTTTCAT; prostate, CCAGG.
TAG_TISSUE=osteoblast
TAG_LIB=UI-R-EP0
TAG_SEQ=AAGATATCAA"

ORIGIN

Alignment Scores: 4.36e-31 Length: 693
Pred. No.: 364.00 Matches: 70
Score: 96.05% Conservative: 3
Percent Similarity: 92.11% Mismatches: 3
Best Local Similarity: 93.09% Indels: 0
Query Match: 13 Gaps: 0
DB:

US-10-029-020-14_COPY_2650_2725 (1-76) x BQ210132 (1-693)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 470 TCCCAATCAACACCATGCTCGTGGCAGACACTAGACGCTACACTGACATCCAGCTCAG 411
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluLysAlaArg 40
Db 410 TACAGGCGACTGTGCTGACACCCGCTATGGGACACCGTGGACGAGGAAAGGTACGG 351
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 350 GTGCTGGAGCTGGCCAGGAGCGCTGTACGCCAGCGCTGGCCCCGAGACGACAGCAGA 291
QY 61 LeuArgGluGluGluGluGluArgAlaTrpThrGluGlyGluLys 76
Db 290 CTGCGGAGAGGAGAGAGGCGCTCGCGGCTTGACACACGCGGAGAGAAG 243

RESULT 10

Al176158/c 514 bp mRNA linear EST 20-JAN-1999
LOCUS Al176158
DEFINITION EST2159739 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone
ROVBL55.3' end, mRNA sequence.

ACCESSION Al176158
VERSION Al176158.1 GI:3726796
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.

REFERENCE 1 (bases 1 to 514)
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kertlavage,A.R. and Adams,M.D.
TITLE Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529

Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.

FEATURES

source Location/Qualifiers

1..514
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="ATCC (inhost):2035740"
/db_xref="taxon:10118"
/clone="ROVBL55"
/clone_lib="Normalized rat ovary, Bento Soares"
/note="Organ: ovary; Vector: pMT3Pac; Site_1: EcoRI;
Site_2: NotI"

ORIGIN

Alignment Scores:
Pred. No.: 1,35e-30 Length: 514
Score: 358.00 Matches: 69
Percent Similarity: 94.74% Conservative: 3
Best Local Similarity: 90.73% Mismatches: 4
Query Match: 91.56% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x A1176158 (1-514)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 443 TCCCAATCAACCCCATGCTGGTGGCAGGACTAGACGCTACACTGCATCCAGCTGCAG 384
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 383 TACAGGGCACTGTGCTGAACACCGCTATGGCAACCGTGGACGAGGAAAAGTACGG 324
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGlnGlnArg 60
Db 323 GTGCTGGAGCTGGCCAGCAGGAGGCTGTACGCCAGGCTGGCCCGCAGACGACGAGA 264
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 263 CTGGCGGAAGGGAAGAGGGCTTGGGGCTGGACAGCGGGGAGAAG 216

RESULT 11
AV589580/c
LOCUS AV589580 Bos taurus brain fetus Bos taurus cDNA clone E1BR007C10
DEFINITION 3', mRNA sequence.
ACCESSION AV589580
VERSION AV589580.1 GI:9700573
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 424)
Takaue, A., Harotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H.
and Sugimoto, Y.
Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs

JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)
MEDLINE 21570554
PUBMED 11713328
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cococn.ocn.ne.jp

Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers

FEATURES

source

1..424
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1BR007C10"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_lib="Bos taurus brain fetus"
/note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"

ORIGIN

Alignment Scores:
Pred. No.: 3.78e-30 Length: 424
Score: 353.00 Matches: 68
Percent Similarity: 94.67% Conservative: 3
Best Local Similarity: 90.67% Mismatches: 4
Query Match: 90.28% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AV589580 (1-424)

QY 2 GlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGlnTyr 21
Db 422 CAGATCAACCCCGATGCTTACGCGAGGACTAGACGCTACNCGACATCCAGCTCCAGTAT 363
QY 22 GlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArgVal 41
Db 362 GGGCTCTGTGCTTGAACACCGCTACGGGACCACTCTGNACGAGAAGAGCGCGGGTG 303
QY 42 LeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArgLeu 61
Db 302 CTGCAGCTGGCCGCGCAGAGCGCTGCGCCAGGCTGGGCCCTGGAGCAGCGACTT 243
QY 62 ArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 242 CGGACCGGAGGAGGGCGCTCGGGCGCTGGACCGAGGCGGAGAAG 198

RESULT 12

AV169909/c

LOCUS AV169909 748 bp mRNA linear EST 29-JAN-2001
DEFINITION AV169909 Ol-br-ad cDNA Oryzias latipes cDNA clone br5373, mRNA
sequence.

ACCESSION AV169909

VERSION AV169909.1 GI:12591978

KEYWORDS EST.

SOURCE Oryzias latipes (Japanese medaka)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

REFERENCE 1 (bases 1 to 748)

Mita, K., Ishikawa, Y. and Yamauchi, M.

Establishment of cDNA database of medaka, Oryzias latipes

Unpublished (2001)

CONTACT Mita K

Genome Research Group

National Institute of Agrobiological Sciences

Osashi 1-2, Tsukuba, Ibaraki 305-8634, Japan

Email: kmita@nias.affrc.go.jp

Method: uni-directional sequence direction: sequenced from T3 primer
(5' -> 3').

FEATURES

source

1..748 Location/Qualifiers

/organism="Oryzias latipes"

/mol_type="mRNA"

/strain="HNT"

/db_xref="taxon:8090"

/clone="br5373"

/sex="female/male mixed"

/tissue_type="brain"

low stringency conditions."

/dev_stage="adult"
/clone_lib="OI-br-ad cdna"

ORIGIN

Alignment Scores:

Pred. No.: 8,99e-29 Length: 748
Score: 344.00 Matches: 65
Percent Similarity: 93.4% Conservatives: 6
Best Local Similarity: 93.4% Mismatches: 5
Query Match: 85.53% Indels: 5
DB: 87.98% Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AU169909 (1-748)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 230 TCTAGGTCACACCGCTGTTCACCGCAGGACTAGCGCATCACAGACATCCAGCTGCAG 171
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgGlyThrThrLeuAspGluGluLysAlaArg 40
Db 170 TATGTGCGCTGTGCTGAACACCGCTACCGCAGCGCTGCGCAGGAGGAGGCGCG 111
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 110 GTTCTGAGCTGCTCGCAGAGGCTGTGACGAGGCTTGGCTCGGAGGCGCAGAGA 51
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 50 CTGAGGAGCGGAGGAGGCTCCCGGCGCTGACTGAGGAGGAGAG 3

RESULT 13

AW580211/c
LOCUS 507 bp mRNA linear EST 16-MAR-2000
DEFINITION RC4-HT0411-080100-013-a03 HT0411 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW580211
VERSION AW580211.1 GI:7255260
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 507)
HCGP <http://www.ludwig.org.br/ORESTES>.
The FAPESP/LiCR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LiCR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC4&t2=RC4-HT0411-080100-013-a03&t3=2000-01-08&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 27
High quality sequence stop: 237.

FEATURES

source
1..507
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0411"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under

ORIGIN

Alignment Scores:

Pred. No.: 5,36e-28 Length: 507
Score: 335.00 Matches: 73
Percent Similarity: 93.59% Conservatives: 0
Best Local Similarity: 93.59% Mismatches: 2
Query Match: 85.68% Indels: 3
DB: 10 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AW580211 (1-507)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 380 TCCAGATCAACACCAAGTACTTATGCGCAAGGACTAGACGCTACACAGACATCCAGCTCC 321
QY 20 LntyrGlyAlaLeuCysLeuAsnThrArgGlyThrThrLeuAspGluGluLysAla 40
Db 320 TGTACGGGGGCACTGTGCTTGAACACACGCTACGGGACACGCTTGGATGAGGAGAGGCAC 261
QY 40 LntyrGlyAlaLeuCysLeuAsnThrArgGlyThrThrLeuAspGluGluLysAla 59
Db 260 GGGTCTCTGAGCTGCTCGGAGAGAGCGCTGCGCAAGCGTGGCGCCGCGAGCATCAG 201
QY 60 ArgLeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGlu 75
Db 200 AGACTCCGGGAAGGGAGGAGGAGGACTCGCGGCTGCACAGAGGGGGAG 153

RESULT 14

AL045769/c
LOCUS 713 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp434F206.s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045769
VERSION AL045769.1 GI:5433881
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 713)
Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Koehler, et al.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charité,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
r1 sequence also available.
This clone (DKFZp434F206) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..713
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="DKFZp434F206"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Alignment Scores: 2.72e-25 Length: 713
Pred. No.:

Score: 313.00 Matches: 68
Percent Similarity: 92.00% Conservative: 1
Best Local Similarity: 90.67% Mismatches: 6
Query Match: 80.05% Indels: 2
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AL045769 (1-713)

QY 2 GlnlleAsnThrValLeuAsnGlyArgThrArgGlyThrArgPileGlnLeuGlnTyr 21
Db 652 CAGATCAACCA-GTACTTAATGGCAGGACTAGAGCTACACAGACATCCAGCTCCAGTAC 594
QY 22 GlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLeuAlaArgVal 41
Db 593 GGGGCCCTGTGCTTGAACCCCGCTACGGGCAAGCTTGGATTAGGAGAGCC-CGGGTC 535
QY 42 LeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArgLeu 61
Db 534 CTGAACTGGCCCGGCAAAACCGTGGCAAGCTGGGCGGCGAGCAGCAGAGACTG 475
QY 62 ArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 474 CGGGAAGGGAGGAGGAGGCTTGGGCGCTTGGACAGAGGGGAGAG 430

RESULT 15
BB161536
LOCUS
DEFINITION
BB161536 RIKEN full-length enriched, 16 days neonate thymus Mus
musculus cDNA clone AL30060L24 3' similar to AB025410 Mus musculus
mRNA for Ten-m1, mRNA sequence.
BB161536
BB161536.2 GI:16268632
EST.
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 608)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sugabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
On Jun 29, 2000 this sequence version replaced gi:8817476.
Contact: Yoshinhide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayate, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Matahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 128.916 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_2650_2725

Perfect score: 391
Sequence: 1 SQINTVNGRTRYDQLQ.....EQQLREGEGELRAWTEGK 76

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N Geneseq 29Jan04 -QWTF=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -FRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptt -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CGN 1.1.1868 @runat_06082004_112215_29265 -NCFU=6 -ICFU=3
-NO MWAP -LARGSEQURY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 391 | 100.0 | 8354 | 6 ABS52100 | Abs52100 Human TEN |
| 2 | 391 | 100.0 | 8438 | 6 ABN85378 | Abn85378 Human NOV |
| 3 | 391 | 100.0 | 8645 | 6 ABS78652 | Abs78652 Human CDN |
| 4 | 311 | 79.5 | 3312 | 4 AAK52812 | Aak52812 Human pol |
| 5 | 311 | 79.5 | 13202 | 4 AAK51828 | Aak51828 Human pol |
| 6 | 308 | 78.8 | 12879 | 6 ABK92230 | Abk92230 Prostate |
| 7 | 296 | 75.7 | 2157 | 4 AAK14096 | Aak14096 Human CDN |
| 8 | 296 | 75.7 | 3270 | 4 AAH14671 | Aah14671 Human CDN |

| | | | | | |
|----|-------|------|-------|-------------|--------------------|
| 9 | 296 | 75.7 | 3614 | 4 AAH14183 | Aah14183 Human CDN |
| 10 | 296 | 75.7 | 8473 | 6 ABQ82345 | Abq82345 Human NOV |
| 11 | 296 | 75.7 | 8487 | 6 ABQ82346 | Abq82346 Human NOV |
| 12 | 296 | 75.7 | 8645 | 6 ABQ82344 | Abq82344 Human NOV |
| 13 | 296 | 75.7 | 8675 | 6 ABQ82343 | Abq82343 Human NOV |
| 14 | 246.5 | 63.0 | 9058 | 7 ACC72051 | Acc72051 BC02050A |
| 15 | 246.5 | 63.0 | 9695 | 7 ACC72052 | Acc72052 BC02050B |
| 16 | 246.5 | 63.0 | 9729 | 5 AAS14089 | Aas14089 Human FCT |
| 17 | 246.5 | 63.0 | 9729 | 5 AAS14089 | Aas14089 Human FCT |
| 18 | 246.5 | 63.0 | 9826 | 5 AAS14085 | Aas14085 Human FCT |
| 19 | 246.5 | 63.0 | 9826 | 5 AD832023 | Ad832023 Human FCT |
| 20 | 135.5 | 34.7 | 540 | 4 AAH10637 | Aah10637 Human CDN |
| 21 | 124 | 31.7 | 3910 | 4 ABLO4849 | Ablo4849 Drosophil |
| 22 | 124 | 31.7 | 6318 | 4 ABLO4848 | Ablo4848 Drosophil |
| 23 | 117.5 | 30.1 | 464 | 8 ACH40709 | Ach40709 Human foe |
| 24 | 103 | 26.3 | 10242 | 4 ABL29075 | Ab129075 Drosophil |
| 25 | 103 | 26.3 | 17131 | 4 ABL29074 | Ab129074 Drosophil |
| 26 | 90 | 23.0 | 65 | 6 ABN55563 | Abn55563 Mouse spl |
| 27 | 77.5 | 19.8 | 372 | 3 AAC08271 | Aac08271 Human sec |
| 28 | 77.5 | 19.8 | 458 | 3 AAC01764 | Aac01764 Human sec |
| 29 | 77.5 | 19.8 | 2015 | 10 ADE77097 | Ade77097 Human CDN |
| 30 | 75.5 | 19.2 | 992 | 6 ABK34678 | Abk34678 Human CDN |
| 31 | 75 | 19.2 | 6457 | 6 ABN96892 | Abn96892 Gene #339 |
| 32 | 73.5 | 18.8 | 285 | 6 ABL69281 | Ab169281 Prostate |
| 33 | 73.5 | 18.8 | 445 | 6 ABT04001 | Abt04001 Human ova |
| 34 | 73.5 | 18.8 | 791 | 6 ABK35440 | Abk35440 Human CDN |
| 35 | 73.5 | 18.8 | 14113 | 3 AAA39651 | Aaa39651 Fusion co |
| 36 | 73.5 | 18.8 | 14113 | 3 AAA90096 | Aaa90096 PR-1 prom |
| 37 | 73.5 | 18.8 | 14113 | 4 AAH25853 | Aah25853 PR-1 prom |
| 38 | 73 | 18.7 | 60 | 6 ABN43073 | Abn43073 Human spl |
| 39 | 72 | 18.4 | 15231 | 6 ABK63598 | Abk63598 Rat sequ |
| 40 | 72 | 18.4 | 15231 | 7 ABT42078 | Abt42078 Toxicity |
| 41 | 72 | 18.4 | 15231 | 9 ADB59205 | Ad59205 Toxicity |
| 42 | 72 | 18.4 | 15231 | 9 ADB53026 | Ad53026 Primary r |
| 43 | 70.5 | 18.0 | 300 | 2 AAZ13121 | Aaz13121 Human gen |
| 44 | 70.5 | 18.0 | 761 | 6 ABQ60799 | Abq60799 Human mgn |
| 45 | 70.5 | 18.0 | 775 | 2 AAZ15180 | Aaz15180 Human gen |

ALIGNMENTS

RESULT 1
ABS52100
ID ABS52100 standard; DNA; 8354 BP.
XX
AC ABS52100;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human TEN-M4-like gene.

Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis; cell signal processing; metabolic pathway modulation; metabolic disorder; obesity; diabetes; infectious disease; neurodegenerative disorder; cancer; Alzheimer's disease; Parkinson's disease; immune disorder; cancer; haematopoietic disorder; cirrhosis; pancreatitis; learning defect; memory defect; infertility; congenital heart defect; hair growth; pigmentation disorder; endocrine disorder; reproductive; neurological disease; gastro-intestinal disease; endocrine disorder; respiratory disease; health; bone marrow transplantation; endocrine disease; allergic inflammation; nephrological disorder; urinary system disorder; age-related disorder; neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4; adipocyte complement-related Clq tumour necrosis factor; out at first; beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter; type Ia membrane sushi-containing domain; butyrophilin; single nucleotide polymorphism.

Homo sapiens.

Key Location/Qualifiers
FT variation replace(117,G)
FT /*tag= a

variation /standard_name= "Single nucleotide polymorphism (SNP)"
 replace(225,C)
 /*tag= b
 /*standard_name= "Single nucleotide polymorphism (SNP)"
 replace(260,A)
 /*tag= c
 /*standard_name= "Single nucleotide polymorphism (SNP)"

WO200257453-A2.

25-JUL-2002.

19-DEC-2001; 2001WO-US050331.

19-DEC-2000; 2000US-0256704P.

20-DEC-2000; 2000US-0257314P.

02-MAY-2001; 2001US-0288153P.

29-MAY-2001; 2001US-0294075P.

24-JUL-2001; 2001US-0307506P.

10-AUG-2001; 2001US-0311590P.

10-AUG-2001; 2001US-0311613P.

29-AUG-2001; 2001US-0315617P.

14-SEP-2001; 2001US-0322358P.

(CURA-) CURAGEN CORP.

Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R, Stone DJ, Anderson D, Shinkets RA, Burgess CE, Zerhusen BD, Liu X, Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W, WPI; 2002-590744/63.
 P-PSDB; ABG70388.

Novel isolated NOVX polypeptide useful for treating cardiomyopathy, atherosclerosis, metabolic disorders, diabetes, obesity, infectious disease, anorexia, neurodegenerative disorders, Alzheimer's disease or cancer.

Claim 8; Page 50-52; 318pp; English.
 The present invention relates to new NOVX polypeptides. The invention is useful for treating or preventing a NOVX-associated disorder such as cardiomyopathy or atherosclerosis, where the disorder is related to cell signal processing and metabolic pathway modulation in a subject, preferably human. The invention is also useful for treating metabolic disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disease), immune disorders, haematopoietic disorders and various cancers. The molecules of the invention are also useful for treating or preventing cirrhosis, pancreatitis, learning and memory defects, infertility, congenital heart defects, acne, hair growth, pigmentation disorders, endocrine disorders, respiratory disease, gastro-intestinal diseases, reproductive health, neurological diseases, bone marrow transplantation, endocrine diseases, allergy and inflammation, nephrological disorders, urinary system disorders, neuropsychiatric disorders and age-related disorders. The present nucleic acid sequence represents a NOVX gene. This sequence encodes a NOVX protein of the invention

Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.53e-39 Length: 8354
 Score: 391.00 Matches: 76
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 6

US-10-029-020-14_COPY_2650_2725 (1-76) x ABS52100 (1-8354)

1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
 DB 7982 TCCAGATCAACACAGTACTTAAATGGCAGACTAGACGCTACACAGATCCAGTCCAG 8041

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
 DB 8042 TACGGGGCACTGTGCTTGAACACACGCTACGGGACACGTTGGATGAGGAGGACCGG 8101
 QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrrAlaArgGluGlnGlnArg 60
 DB 8102 GTCTCTGGAGCTGGCCCGGCAGAGAGCCGTGGCCCAAGCGTGGCCCGCCGAGCAGAGA 8161
 QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrrThrGluGlyGluLys 76
 DB 8162 CTGCGGGAAGGGGAGGAGGCTTGGCGGCTGACAGAGGGGGGAGAG 8209
 RESULT 2
 ABN85378
 ID ABN85378 standard; DNA; 8438 BP.
 XX
 AC ABN85378;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human NOV1, TEN-M4 like protein, coding sequence.
 XX
 KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
 KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
 KW Antithrombotic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
 KW Antibacterial; Vitucide; Antiparasitic; Relaxant; Anticonvulsant;
 KW Gene Therapy; NOV; cancer; heart disease; inflammation;
 KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
 KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
 KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
 KW TEN-M4 like protein; chromosome 11; gene; ds.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 CDS 4..8395
 FT /*tag= a
 FT /trans except= (pos: 1138..1147,aa:Met)
 FT /product= "NOV1 protein"

MO20025704-A2.

18-JUL-2002.

09-JAN-2002; 2002WO-US000554.

09-JAN-2001; 2001US-0260417P.

10-JAN-2001; 2001US-0260831P.

28-FEB-2001; 2001US-0272338P.

09-MAR-2001; 2001US-0274876P.

18-APR-2001; 2001US-0284704P.

(CURA-) CURAGEN CORP.

padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA, Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM, Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM, Alsbrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K, Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G, Gunther E, Stone DJ;
 WPI; 2002-590674/63.
 P-PSDB; ABB98401.

NOVX polypeptides and encoding polynucleotides, useful for preventing or treating NOVX-associated disorders e.g. cancer, inflammation, or Alzheimer's disease, and in chromosome mapping, tissue typing or pharmacogenomics.

Claim 9; Page 8-9; 358pp; English.

The present sequence is a coding sequence for a NOV protein. The NOV

CC

CC proteins and coding sequences are useful for treating or preventing NOV-
 CC associated disorders or in the manufacture of a medicament for treating NOV-
 CC the disorders, such as cancer, heart disease, inflammation, autoimmune
 CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
 CC IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
 CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
 CC and other wasting disorders associated with chronic diseases. NOV1 is a
 CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
 XX
 SQ Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.59e-39 Length: 8438
 Score: 391.00 Matches: 76
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABN85378 (1-8438)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
 DB TCCGAGATCAACACAGTACTTATGGCAGAGCTAGACGCTACACAGATCCAGCTCCAG 8092
 QY 21 TyrGlyAlaLeuGlyLeuAsnThrArgTyrGlyThrThrLeuAspGluGluysAlaArg 40
 DB TACGGGGCACTGTGCTTGAACACACAGCTACGGGACAACTTGATGAGAGAGGACGCG 8152
 QY 41 ValLeuGluLeuAlaArgGlnAlaValArgGlnAlaTrrPalaArgGluGlnGlnArg 60
 DB GTCCCTGGAGCTGGCCCGCAGAGAGCGTGGCCCAAGCGTGGCCCGCAGCAGCAGAGA 8212
 QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrrThrGluGlyGluys 76
 DB CTGCGGAGAGGAGAGAGGAGGCTGGCGGCTGGAGAGGAGGAGAGGAGAG 8260

RESULT 3

AB578652
 ID ABS78652 standard; cDNA; 8645 BP.
 XX
 AC ABS78652;
 XX
 DT 16-DEC-2002 (first entry)
 XX
 DE Human cDNA encoding CGDD10, INCYTE 7488573CB1.

XX Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
 KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
 KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
 KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
 KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
 KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
 KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
 KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
 KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
 KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
 KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
 KW rheumatoid arthritis.

OS Homo sapiens.

XX

XX WO200272830-A2.

PN

PD 19-SEP-2002.

XX

PF 08-FEB-2002; 2002WO-US0003715.

XX

PR 09-FEB-2001; 2001US-0268111P.

PR 23-FEB-2001; 2001US-0271175P.

PR 08-MAR-2001; 2001US-0274503P.

PR 09-MAR-2001; 2001US-0274552P.

XX

(INCYTE) INCYTE GENOMICS INC.

Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
 Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
 Lu DAM, Richardson TW, Tran UK, Khare R, Wallia NK;
 WPI; 2002-723356/78.
 DR F-PSDB; ABG97359.

XX New human proteins associated with cell growth, differentiation and
 PT death, useful for diagnosing, treating or preventing autoimmune or
 PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
 PT atherosclerosis or hepatitis.

Claim 5; Page 175-178; 181pp; English.

CC The invention relates to an isolated polypeptide comprising CGDD1-12
 CC (cell growth, differentiation and death), a naturally occurring amino
 CC acid sequence at least 90% identical to CGDD, a biologically active
 CC fragment or an immunogenic fragment. Also included are the
 CC polynucleotides encoding CGDD1-12. a recombinant polynucleotide
 CC comprising a promoter sequence operably linked to the CGDD
 CC polynucleotides, a cell transformed with the recombinant polynucleotide
 CC a transgenic organism comprising the recombinant polynucleotide, an anti-
 CC CGDD antibody, screening for compounds which bind to/modulate or are
 CC antagonists of CGDD or alter the expression of CGDD polynucleotide and a
 CC CGDD polynucleotide microarray. The polypeptides, polynucleotides, and
 CC agonists and antagonists are useful for diagnosing, treating or
 CC preventing disorders associated with aberrant expression of CGDD,
 CC particularly cell proliferation (e.g. arteriosclerosis, atherosclerosis,
 CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
 CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
 CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
 CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
 CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
 CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
 CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
 CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
 CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
 CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
 CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
 CC infections. They are also useful in the assessment of the effects of
 CC exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of proteins associated with CGDD. The present sequence encodes
 CC a CGDD protein

SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 4.73e-39 Length: 8645
 Score: 391.00 Matches: 76
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABS78652 (1-8645)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
 DB TCCGAGATCAACACAGTACTTATGGCAGAGCTAGACGCTACACAGATCCAGCTCCAG 8090
 QY 21 TyrGlyAlaLeuGlyLeuAsnThrArgTyrGlyThrThrLeuAspGluGluysAlaArg 40
 DB TACGGGGCACTGTGCTTGAACACACAGCTACGGGACAACTTGATGAGAGAGGACCGG 8150
 QY 41 ValLeuGluLeuAlaArgGlnAlaValArgGlnAlaTrrPalaArgGluGlnGlnArg 60
 DB GTCCCTGGAGCTGGCCCGCAGAGAGCGTGGCCCAAGCGTGGCCCGCAGCAGCAGAGA 8210
 QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrrThrGluGlyGluys 76
 DB CTGCGGAGAGGAGAGAGGAGGCTGGCGGCTGGAGAGGAGGAGAGGAGAG 8258

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
 Db 3017 TCCAGATGACTTCTGTGTTGAATGGGAGGACTAGACGGTTTGGAGATATTTCAGCTCCAG 3076
 QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
 Db 3077 CATGGAGCCCTGTGCTTCAACATCCGGTATGGGACACTGTGGAAGAGAAAGATCATC 3136
 QY 41 ValLeuGluLeuAlaAlaArgGlnArgAlaValArgGlnAlaThrAlaArgGluGlnArg 60
 Db 3137 GTGTTGGAGATGTCAGACAGCGCGCTAGTGGCCCGCTGGACTTAAGACACAAGAGG 3196
 QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaThrThrGluGlyGluLys 76
 Db 3197 CTGCAAGAGGGGAAAGAGGGGATTAGGGCATGGACAGAGAGGGGAAAG 3244
 RESULT 5
 AAK51828
 ID AAK51828 standard; cDNA; 13202 BP.
 XX
 AC AAK51828;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 373.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI; 2001-476283/51.
 DR P-PSDB; AAM79679.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 PS Claim 1; Page 4651-4652; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 3312 BP; 928 A; 750 C; 739 G; 895 T; 0 U; 0 Other;

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
 Db 3017 TCCAGATGACTTCTGTGTTGAATGGGAGGACTAGACGGTTTGGAGATATTTCAGCTCCAG 3076
 QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
 Db 3077 CATGGAGCCCTGTGCTTCAACATCCGGTATGGGACACTGTGGAAGAGAAAGATCATC 3136
 QY 41 ValLeuGluLeuAlaAlaArgGlnArgAlaValArgGlnAlaThrAlaArgGluGlnArg 60
 Db 3137 GTGTTGGAGATGTCAGACAGCGCGCTAGTGGCCCGCTGGACTTAAGACACAAGAGG 3196
 QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaThrThrGluGlyGluLys 76
 Db 3197 CTGCAAGAGGGGAAAGAGGGGATTAGGGCATGGACAGAGAGGGGAAAG 3244
 RESULT 5
 AAK51828
 ID AAK51828 standard; cDNA; 13202 BP.
 XX
 AC AAK51828;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 2341.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI; 2001-476283/51.
 DR P-PSDB; AAM79679.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 PS Claim 1; Page 4651-4652; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 3312 BP; 928 A; 750 C; 739 G; 895 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.23e-29 Length: 3312
 Score: 311.00 Matches: 57
 Percent Similarity: 89.47% Conservative: 11
 Best Local Similarity: 75.00% Mismatches: 8
 Query Match: 79.54% Indels: 0
 DB: 4 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AAK52812 (1-3312)

CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.26e-28 Length: 13202
 Score: 311.00 Matches: 57
 Percent Similarity: 89.47% Conservativity: 11
 Best Local Similarity: 75.00% Mismatches: 8
 Query Match: 79.54% Indels: 0
 DB: 4 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AAK51828 (1-13202)

Qy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
 Db 8207 TCCGAGATGACTTCTGTGTGATGGAGGACTAGACGGTTTGCAGATATTCAGCTCCAG 8266
 Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluysAlaArg 40
 Db 8267 CATGGAGCCCTGTGCTTCAACATCCGGTATGGACAACTGTGCAAGAGGAAAGATCAC 8326
 Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
 Db 8327 GTGTGGAGATTGCCAGACAGCGCCAGTGGCCCGAGCTGGACTAAGGAAACAAGAAGG 8386
 Qy 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
 Db 8387 CTGCAAGGGGGAGAGGGGATTAGGGCATGGAGAGAGGAGAAAG 8434

RESULT 6

ABK92230
 ID ABK92230 standard; DNA; 12879 BP.

AC ABK92230;

XX
 DT 15-AUG-2002 (first entry)

XX Prostate cancer-associated DNA sequence #116.

XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
 gene therapy; gene; ds.

XX Mammalia.

XX WO200230268-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032045.

XX 13-OCT-2000; 2000US-00687576.

XX 08-DEC-2000; 2000US-00733288.

XX 08-DEC-2000; 2000US-00733742.

XX 24-JAN-2001; 2001US-0263957P.

XX 16-MAR-2001; 2001US-0276791P.

XX 16-MAR-2001; 2001US-0276888P.

XX 06-APR-2001; 2001US-0281922P.

XX 24-APR-2001; 2001US-0286214P.

XX 30-APR-2001; 2001US-00847046.

XX 04-MAY-2001; 2001US-0288589P.

XX (BOSB-) EOS BIOTECHNOLOGY INC.

XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX WPI; 2002-471335/50.

XX P-PSDB; ABG61913.

XX Detecting a prostate cancer-associated transcript in a cell in a patient,
 PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
 PT by determining if prostate cancer-associated genes are expressed in a

PT prostate tissue.

XX Claim 22; Page 394-397; 436pp; English.

XX The present invention relates to methods of detecting a prostate cancer-

CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with prostate cancer-
 CC associated polynucleotides (designated PC genes) that selectively
 CC hybridise to a sequence that is at least 80% identical to them. The
 CC prostate cancer-associated polynucleotide sequences are differentially
 CC expressed in prostate tumour tissue or in prostate cancer and are derived
 CC from the tissues of various organisms such as humans or other mammals
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
 CC associated genes are useful for diagnosing or treating prostate cancer,
 CC as well as for identifying modulators of prostate cancer or agents that
 CC inhibit prostate cancer. The nucleic acid sequences are particularly
 CC useful in gene therapy, as a vaccine or in antisense applications.
 CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
 CC sequences

SQ Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.96e-28 Length: 12879
 Score: 308.00 Matches: 56
 Percent Similarity: 89.47% Conservativity: 12
 Best Local Similarity: 73.68% Mismatches: 8
 Query Match: 78.77% Indels: 0
 DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABK92230 (1-12879)

Qy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
 Db 7880 TCCGAGATGACTTCTGTGTGATGGAGGACTAGACGGTTTGCAGATATTCAGCTCCAG 7939

Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluysAlaArg 40
 Db 7940 CATGGAGCCCTGTGCTTCAACATCCGGTATGGACAACTGTGCAAGAGGAAAGATCAC 7999

Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
 Db 8000 GTGTGGAGATTGCCAGACAGCGCCAGTGGCCCGAGCTGGACTAAGGAAACAAGAAGG 8059

Qy 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
 Db 8060 CTGCAAGAGGGGAGAGGGGATTAGGGCATGGACAGAGGGGAAAG 8107

RESULT 7

AAH14096

ID AAH14096 standard; cDNA; 2157 BP.

XX AC AAH14096;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:11265.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EF1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

Mon Aug 16 09:01:14 2004

us-10-029-020-14_copy_2650_2725.p2n.rng

PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 XX Claim 8; SEQ ID NO 11265; 2537pp + Sequence Listing; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises at least 15 nucleotides and the combination of
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 XX Sequence 2157 BP; 544 A; 560 C; 591 G; 462 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 1.06e-27 Length: 2157
 Score: 296.00 Matches: 54
 Percent Similarity: 86.84% Conservative: 12
 Best Local Similarity: 71.05% Mismatches: 10
 Query Match: 75.70% Indels: 0
 DB: 4 Gaps: 0
 US-10-029-020-14_COPY_2650_2725 (1-76) x AAH14096 (1-2157)
 QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
 Db 1550 TCGCAGTCCACCCAGCGTGGTGTGAACCGGACGAGCGGAGTTCGGCGAGTGCAG 1609
 QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
 Db 1610 TTCGGCGCGTGGCGCTGCACGTGCTACGGCATGACCTTCGACGAGGAGCGCGC 1659
 QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
 Db 1670 ATCTGGAGAGCGCGCGAGCGCGCTCGCCGGCGCTGGCGCGAGCAGCGC 1729
 QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
 Db 1730 GTGCGCGAGCGCGAGGCGCGCGCTCTTGGACGCGAGGCGGAGAAG 1777
 RESULT 8
 AAH14671
 ID AAH14671 standard; cDNA; 3270 BP.
 XX
 XX AAH14671;
 AC
 XX

DT 26-JUN-2001 (first entry)
 DE Human cDNA sequence SEQ ID NO:12354.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX Homo sapiens.
 OS
 XX EP1074617-A2.
 PN
 XX 07-FEB-2001.
 PD
 XX 28-JUL-2000; 2000EP-00116126.
 PF
 XX 29-JUL-1999; 95JP-00248036.
 PR 27-AUG-1999; 95JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 PI WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 XX Claim 8; SEQ ID NO 12354; 2537pp + Sequence Listing; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises at least 15 nucleotides and the combination of
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 XX Sequence 3270 BP; 867 A; 821 C; 852 G; 730 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 1.79e-27 Length: 3270
 Score: 296.00 Matches: 54
 Percent Similarity: 86.84% Conservative: 12
 Best Local Similarity: 71.05% Mismatches: 10
 Query Match: 75.70% Indels: 0
 DB: 4 Gaps: 0
 US-10-029-020-14_COPY_2650_2725 (1-76) x AAH14671 (1-3270)
 QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
 Db 2580 TCGCAGTCCACCCAGCGTGGTGTGAACCGGACGAGCGGAGTTCGGCGAGTGCAG 2639


```

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluysAlaArg 40
Db 2640 TTCGGCGGCTGGCGTGCACGTGGCTAGCGATGACCTGGACAGAGAGGCGCGC 2699
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrrAlaArgGluGlnArg 60
Db 2700 ATCTGTGAGCAGCGCGGCGCGCTGCGCGCGCTGCGCGCGCGCGAGCAGCGC 2759
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrrThrGluGlyGluys 76
Db 2760 GTGCGGCGAGCGAGGAGGCGCGCTCTGACGAGGCGGCGAGAG 2807

RESULT 9
ID AAH14183
XX AAH14183 standard; cDNA; 3614 BP.
AC AAH14183;
DT 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:11430.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW Homo sapiens.
OS
XX EP1074617-A2.
FN 07-FEB-2001.
PD
XX 28-JUL-2000; 2000EP-00116126.
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ieshi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX Claim 8; SEQ ID NO 11430; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dt primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the

```

```

CC present invention
XX
SQ Sequence 3614 BP; 967 A; 898 C; 928 G; 821 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,03e-27 Length: 3614
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservative: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AAH14183 (1-3614)
QY 1 SerGlnIleAsnThrValLeuAsnGlyArgTyrGlyThrArgTyrThrAspGlnLeuGln 20
Db 2924 TCGCAGTCCACACCGTGTGTAACGCGAGGACGCCAGGTTCGCGGAGTGAGATGCAG 2983
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluysAlaArg 40
Db 2984 TTCGGCGGCTGGCGCTGACGTGCGCTACGGCATGACCTGGAGAGAGGCGCGC 3043
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrrAlaArgGluGlnArg 60
Db 3044 ATCTGTGAGCAGCGCGGCGCGCTGCGCGCGCTGCGCGCGCGCGAGCAGCGC 3103
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrrThrGluGlyGluys 76
Db 3104 GTGCGGCGAGCGGAGGCGCGCTCTGACGAGGCGGAGAG 3151

RESULT 10
ABQ82345
ID ABQ82345 standard; cDNA; 8473 BP.
XX
AC ABQ82345;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15c encoding cDNA SEQ ID NO:39.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 258..8144
FT /*tag= a
FT /product= "NOV15c"
XX
WO200262999-A2.
XX
PD 15-AUG-2002.
XX
XX 31-DEC-2001; 2001WO-US049976.
XX
XX 29-DEC-2000; 2000US-0258928P.
XX 02-JAN-2001; 2001US-0259415P.
XX 04-JAN-2001; 2001US-0259785P.
XX 20-FEB-2001; 2001US-0269814P.
XX 09-MAR-2001; 2001US-0279863P.
XX 29-MAR-2001; 2001US-0279832P.
XX 29-MAR-2001; 2001US-0279833P.
XX 13-APR-2001; 2001US-0283889P.

```

18-APR-2001; 2001US-0284447P.
25-APR-2001; 2001US-0286683P.
29-MAY-2001; 2001US-0294080P.
16-AUG-2001; 2001US-0312915P.
17-AUG-2001; 2001US-0313325P.
17-SEP-2001; 2001US-0322699P.
26-NOV-2001; 2001US-0333350P.
(CURA-) CURAGEN CORP.
Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
WPI; 2002-732706/79.
P-PSDB; ABP53588.
New NOVX polypeptides and polynucleotides useful for treating NOVX-
associated disorders, such as cancers, neurological disorders, disorders
of vesicular transport, gastrointestinal disorders, and autoimmune
diseases.
Claim 8; Page 119-121; 444pp; English.
The present invention describes novel human proteins designated NOVX,
where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
cytostatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,
antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
protozoacide and antihelminthic activities, and can be used in gene
therapy. The NOVX proteins, nucleotides or antibodies can be used in the
manufacture of a medicament for treating a syndrome associated with a
human disease selected from NOVX-associated disorder, such as cancers
(e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
ischemic cerebrovascular disease, Alzheimer's disease or Pick's
disease), disorders of vesicular transport (e.g. cystic fibrosis,
diabetes mellitus, Grave's disease, or goitre), gastrointestinal
disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
and protozoal infections. The NOVX proteins can be used as immunogens to
produce antibodies and as vaccines. The NOVX nucleotide sequences may be
used in chromosome mapping, identifying individuals from minute
biological samples (tissue typing), and in forensic identification of a
biological sample. The present sequence encodes human NOV15c, which is
located on chromosome 4
XX SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;
Alignment Scores:
Pred. No.: 5, 92e-27 Length: 8473
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservaive: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 6 Gaps:
US-10-029-020-14_COPY_2650_2725 (1-76) x ABQ82345 (1-8473)
Oy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 7782 TCGCAGTCCACCGGTGTGAACGCGAGCGCGAGTTCGCGGACGTGGAGATGACG 7841
Oy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluLysAlaArg 40
Db 7842 TTCGGCGCGCGTGGCCGCTGACGTCACGTCATCGGTCATCCCTGCAGCAGAGCGCGC 7901
Oy 41 ValLeuGluLeuAlaAargGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 7902 ATCCTTGAGCAGCGCGCGAGCGCGCGCTCGCCCGCGCTCGCGCGCGAGCAGCGC 7961

Oy 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 7962 GTGCGGACGCGGAGGAGGCGCGCTCTGCGAGCGGCGGAGAG 8009
RESULT 11
ABQ82346
ID ABQ82346 standard; cDNA; 8487 BP.
XX AC ABQ82346;
XX DT 17-DEC-2002 (first entry)
XX DE Human NOV15d encoding cDNA SEQ ID NO:41.
XX KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;
antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; goitre;
KW Pick's disease; vesicular transport disease; cystic fibrosis; diabetes;
diabetes mellitus; Grave's disease; gastrointestinal disorder; infection;
KW ulcerative colitis; gastric disorder; duodenal disorder; vaccine;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
rheumatoid arthritis; gene; chromosome 4; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
CDS 299..8140
FT /*tag= a
FT /product= "NOV15d"
XX PN WO200262999-A2.
XX PD 15-AUG-2002.
XX PF 31-DEC-2001; 2001WO-US049976.
XX PR 29-DEC-2001; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0283889P.
PR 13-APR-2001; 2001US-0284447P.
PR 18-APR-2001; 2001US-0286683P.
PR 25-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
(CURA-) CURAGEN CORP.
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
WPI; 2002-732706/79.
P-PSDB; ABP53589.
XX DR 2002-732706/79.
XX DR P-PSDB; ABP53589.
XX PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
associated disorders, such as cancers, neurological disorders, disorders
of vesicular transport, gastrointestinal disorders, and autoimmune
diseases.
XX PT Claim 8; Page 123-125; 444pp; English.

XX The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytotatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal,
CC disorders (e.g. ulcerative colitis, or goitre), gastrointestinal,
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15d, which is
CC located on chromosome 4

XX SQ Sequence 8487 BP; 2306 A; 2029 C; 2148 G; 2001 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 5,938-27 Length: 8487
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservativity: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABQ82346 (1-8487)

Qy 1 SerGlnLeuAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleLeuGln 20
Db 7778 TCGCAGTCCACCGGTGGTGAACGGCAGCGCGCGCGTTCGGCGGATGATGAG 7837

Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGlyAlaArg 40
Db 7838 TTCGCGCGCTGGCGCTGCACCTGCTACGCGTACCGTACCGTGGCAGGAGCGCGCG 7897

Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 7898 ATCTGGAGCAGCGCGCGCGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCG 7957

Qy 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluGly 76
Db 7958 GTGCGCAGCGCGAGGAGGCGCGCGCTGCGCGCGCGCGCGCGCGCGAG 8005

RESULT 12
ABQ82344
ID ABQ82344 standard; cDNA; 8645 BP.
XX ABQ82344;
AC ABQ82344;
DT 17-DEC-2002 (first entry)
XX Human NOV15b encoding cDNA SEQ ID NO:37.

XX Human; NOVX; cytotatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; allergic reaction; autoimmune haemolytic anaemia;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;

XX rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 151..8316

XX /*tag= a

XX /product= "NOV15b"

XX WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

XX 29-DEC-2000; 2000US-0258928P.

XX 02-JAN-2001; 2001US-0259415P.

XX 04-JAN-2001; 2001US-0259785P.

XX 20-FEB-2001; 2001US-0269814P.

XX 09-MAR-2001; 2001US-0279863P.

XX 29-MAR-2001; 2001US-0279832P.

XX 13-MAR-2001; 2001US-0279833P.

XX 13-APR-2001; 2001US-0283889P.

XX 18-APR-2001; 2001US-0284447P.

XX 25-APR-2001; 2001US-0286683P.

XX 29-MAY-2001; 2001US-0294080P.

XX 16-AUG-2001; 2001US-0312915P.

XX 17-AUG-2001; 2001US-0313325P.

XX 17-SEP-2001; 2001US-0322699P.

XX 26-NOV-2001; 2001US-0333350P.

XX (CURA-) CURAGEN CORP.

XX Spytek KA, Li L, Wolenc AR, Vernet CM, Eisen A, Liu X;

XX Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;

XX Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;

XX Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;

XX Gunther E, Smithson G, Millet I, Macdougall JR;

XX WPI; 2002-732706/79.

XX P-PSDB; ABP53587.

XX New NOVX polypeptides and polynucleotides useful for treating NOVX-associated disorders, such as cancers, neurological disorders, disorders of vesicular transport, gastrointestinal disorders, and autoimmune diseases.

XX Claim 8; Page 114-117; 444pp; English.

XX The present invention describes novel human proteins designated NOVX, where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective, cytotatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular, antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide, immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide, protozoacide and antihelminthic activities, and can be used in gene therapy. The NOVX proteins, nucleotides or antibodies can be used in the manufacture of a medicament for treating a syndrome associated with a human disease selected from NOVX-associated disorder, such as cancers (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle, ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke, ischaemic cerebrovascular disease, Alzheimer's disease or Pick's disease), disorders of vesicular transport (e.g. cystic fibrosis, diabetes mellitus, Grave's disease, or goitre), gastrointestinal, disorders (e.g. ulcerative colitis, or goitre), autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic and protozoal infections. The NOVX proteins can be used as immunogens to produce antibodies and as vaccines. The NOVX nucleotide sequences may be used in chromosome mapping, identifying individuals from minute biological samples (tissue typing), and in forensic identification of a biological sample. The present sequence encodes human NOV15b, which is located on chromosome 4

SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 6.07e-27 Length: 8645
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservativity: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABQ82344 (1-8645)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 7954 TCGCAGTCCACCACGCGTGTGACGCGCAGCGCGCTGCGGACGTGGAGAGCG 8013
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluLysAlaArg 40
Db 8014 TTCGGCGCGCTGGCGTGCACGTGCGCTACGGCATCACCTGGAGGAGAGCGCGC 8073
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 8074 ATCCTGGAGCGCGCGCGCGCGCTCGCGCGCGCTGCGCGCGCGCGCGCGCGC 8133
QY 61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8134 GTGCGCGCGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGC 8181

RESULT 13
ABQ82343
ID ABQ82343 standard; cDNA; 8675 BP.

AC ABQ82343;
XX
XX
DT 17-DEC-2002 (first entry)

DE Human NOV15a encoding cDNA SEQ ID NO:35.

XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nontropic; antidiabetic; antiinflammatory; fungicide;
KW antihemetic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antinaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

FH CDS 151..8328

FT /*tag= a

FT /product= "NOV15a"

FT /transl_except= (pos:1249..1251,aa:Ser)

XX WO200262959-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

XX 29-DEC-2000; 2000US-0258928P.

XX 02-JAN-2001; 2001US-0259415P.

XX 04-JAN-2001; 2001US-0259785P.

XX 20-FEB-2001; 2001US-0269814P.

XX 09-MAR-2001; 2001US-0279863P.

XX 29-MAR-2001; 2001US-0279832P.

XX 29-MAR-2001; 2001US-0279833P.

XX 13-APR-2001; 2001US-0283889P.

XX 18-APR-2001; 2001US-0284447P.

PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333505P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev F, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog V, Burgess CE, Binger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
DR WPI: 2002-732706/79.
DR P-PSDB; ABP53586.
XX
XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
XX Claim 8; Page 110-112; 444pp; English.

XX The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, nontropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiallergic, virucide,
CC immunosuppressive, antihelminthic, antinaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis). The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15a, which is
CC located on chromosome 4

XX Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 6.09e-27 Length: 8675
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservativity: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABQ82343 (1-8675)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
Db 7966 TCGCAGTCCACCACGCGTGTGACGCGCAGCGCGCTGCGGACGTGGAGATGCG 8025
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluLysAlaArg 40
Db 8026 TTCGGCGCGCTGGCGCTGCACGTGCGCTACGCATGACCTTGGAGAGAGCGCGC 8085
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 8086 ATCCTGGAGCGCGCGCGCGCGCTCGCGCGCGCGCTGCGCGCGCGCGCGCGCGC 8145

```
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8146 GTGCGCGAGCGGAGGAGGCGCGCTCTGACGAGGCGAGAG 8193
RESULT 14
ACC72051
ID ACC72051 standard; DNA; 9058 BP.
XX
AC ACC72051;
DT 08-JUL-2003 (first entry)
DE BC0205A gene #SEQ ID 79.
XX
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX
OS Homo sapiens.
XX
FN WO2003029421-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031287.
XX
PR 03-OCT-2001; 2001US-0326526P.
PR 14-MAY-2002; 2002US-00144194.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
DR WPI; 2003-381623/36.
DR P-PSDB; ABR58317.
XX
PT New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
PS Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.
XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records,
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.31e-20 Length: 9058
Score: 246.50 Matches: 49
Percent Similarity: 75.95% Conservative: 11
Best Local Similarity: 62.03% Mismatches: 16
Query Match: 63.04% Indels: 3
DB: 7 Gaps: 1
US-10-029-020-14_COPY_2650_2725 (1-76) x ACC72051 (1-9058)
QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 7328 TCCAGCCCGAGCGTGTGTCACGCGAGGACTCGAGGTTTCAGACATTGAGTTCGAG 7387
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu 37
Db 7388 TACTCACCGCTGCTGCTCAGCATCCCTATGGCCTCACCCCGACACCTTGGACGAAGAG 7447
QY 38 LysAlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGlu 57
Db 7448 AAGGCCCGCGTCTGTGACCGAGGAGACAGAGGGCCCTGGGCACCGCCTGGGCAAGGAG 7507
QY 58 GlnGlnArgLeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 7508 CAGCAGAAAGCAGGAGCGGAGAGAGGGGCGCCCTGTGGACTGAGGCGGAGAAAG 7564
RESULT 15
ACC72052
ID ACC72052 standard; DNA; 9695 BP.
XX
AC ACC72052;
DT 08-JUL-2003 (first entry)
DE BC0205B gene #SEQ ID 81.
XX
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX
OS Homo sapiens.
XX
FN WO2003029421-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031287.
XX
PR 03-OCT-2001; 2001US-0326526P.
PR 14-MAY-2002; 2002US-00144194.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
DR WPI; 2003-381623/36.
DR P-PSDB; ABR58318.
XX
PT New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.
XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records,
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.43e-20 Length: 9695
Score: 246.50 Matches: 49
Percent Similarity: 75.95% Conservative: 11
Best Local Similarity: 62.03% Mismatches: 16
Query Match: 63.04% Indels: 3
DB: 7 Gaps: 1
```

Mon Aug 16 09:01:14 2004

US-10-029-020-14_COPY_2650_2725 (1-76) x ACC72052 (1-9695)

| | | | |
|----|------|--|------|
| Qy | 1 | SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln | 20 |
| Db | 7965 | TCCGAGCCACGCTGCTCAACGGCAGGACTCGAAGGTTACCGACATTGAGTTCCAG | 8024 |
| Qy | 21 | TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu | 37 |
| Db | 8025 | TACTCCAGCGTCTGCTCAGCATCCGCTATGGCCTCACCCCGACACCCCTGGACGAGAG | 8084 |
| Qy | 38 | LysAlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTyrAlaArgGlu | 57 |
| Db | 8085 | AAGGCCCGCTCTGGACCGAGCAGAGAGGCGCTGGGCGACCGGCTGGGCGCAGGAG | 8144 |
| Qy | 58 | GlnGlnArgLeuArgGluGluGluGluGluGluGluGluGluGluGluGluGluGlu | 76 |
| Db | 8145 | CAGCAGAAAGCCAGGAG | 8201 |

Search completed: August 14, 2004, 02:35:01
Job time : 153.916 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 402.845 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_2400_2600

Perfect score: 1077

Sequence: 1 IGVGGLYDLPLTKLVHGR.....TDIISVANEDGRRVAAILNH 201

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112218_29331/app query.fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100
-NCPU=6 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10029020 -CGN 1_1 2156 -runat_06082004_112218_29331
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=0 -WAIT -DSPBLOCK=100
-FGAPOF=6 -FGAPEXT=7 -YGAPOF=10 -YGAPEXT=0.5 -DELOP=10 -XGAPEXT=0.5
-DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Score | Match Length | ID | Description |
|------------|-------------|--------------|----|-------------|
|------------|-------------|--------------|----|-------------|

| | | | | | | |
|----|-------|-------|---------|----|---------------------|-------------------|
| 1 | 1077 | 100.0 | 8354 | 13 | US-10-383-201-43 | Sequence 43, Appl |
| 2 | 1077 | 100.0 | 8354 | 13 | US-10-029-020-13 | Sequence 13, Appl |
| 3 | 1063 | 98.7 | 8355 | 13 | US-10-383-201-55 | Sequence 55, Appl |
| 4 | 1063 | 98.7 | 8438 | 13 | US-10-042-865-1 | Sequence 1, Appl |
| 5 | 630.5 | 58.5 | 12880 | 16 | US-10-295-027-927 | Sequence 927, App |
| 6 | 566 | 52.6 | 3614 | 13 | US-10-342-887-1743 | Sequence 1743, Ap |
| 7 | 566 | 52.6 | 3614 | 13 | US-10-172-118-1743 | Sequence 1743, Ap |
| 8 | 566 | 52.6 | 8473 | 17 | US-10-038-854-39 | Sequence 39, Appl |
| 9 | 566 | 52.6 | 8487 | 17 | US-10-038-854-41 | Sequence 41, Appl |
| 10 | 566 | 52.6 | 8645 | 17 | US-10-038-854-37 | Sequence 37, Appl |
| 11 | 566 | 52.6 | 8675 | 17 | US-10-038-854-35 | Sequence 35, Appl |
| 12 | 556 | 51.6 | 791 | 13 | US-09-823-245A-85 | Sequence 85, Appl |
| 13 | 536.5 | 49.8 | 8797 | 9 | US-09-808-602-74 | Sequence 74, Appl |
| 14 | 536.5 | 49.8 | 8797 | 9 | US-09-808-602-77 | Sequence 77, Appl |
| 15 | 536.5 | 49.8 | 8797 | 10 | US-09-800-198-62 | Sequence 62, Appl |
| 16 | 536.5 | 49.8 | 8797 | 10 | US-09-800-198-65 | Sequence 65, Appl |
| 17 | 531.5 | 49.4 | 8689 | 9 | US-09-808-602-78 | Sequence 78, Appl |
| 18 | 531.5 | 49.4 | 8689 | 10 | US-09-800-198-66 | Sequence 66, Appl |
| 19 | 529.5 | 49.2 | 3217 | 15 | US-10-198-846-13976 | Sequence 13976, A |
| 20 | 529.5 | 49.2 | 6560 | 9 | US-09-808-602-76 | Sequence 76, Appl |
| 21 | 529.5 | 49.2 | 6560 | 10 | US-09-800-198-64 | Sequence 64, Appl |
| 22 | 529.5 | 49.2 | 9058 | 16 | US-10-144-194A-79 | Sequence 79, Appl |
| 23 | 529.5 | 49.2 | 9695 | 16 | US-10-144-194A-81 | Sequence 81, Appl |
| 24 | 529.5 | 49.2 | 9729 | 9 | US-09-808-602-12 | Sequence 12, Appl |
| 25 | 529.5 | 49.2 | 9729 | 10 | US-09-800-198-12 | Sequence 12, Appl |
| 26 | 529.5 | 49.2 | 9826 | 9 | US-09-808-602-7 | Sequence 7, Appl |
| 27 | 529.5 | 49.2 | 9826 | 10 | US-09-800-198-7 | Sequence 7, Appl |
| 28 | 525.5 | 48.8 | 8409 | 9 | US-09-808-602-79 | Sequence 79, Appl |
| 29 | 525.5 | 48.8 | 8409 | 10 | US-09-800-198-67 | Sequence 67, Appl |
| 30 | 520.5 | 48.3 | 8575 | 13 | US-10-072-012-143 | Sequence 143, App |
| 31 | 194 | 18.0 | 431 | 9 | US-09-954-456-1034 | Sequence 1034, Ap |
| 32 | 188.5 | 17.5 | 1973 | 9 | US-09-864-761-4526 | Sequence 4526, Ap |
| 33 | 131.5 | 12.2 | 399 | 9 | US-09-861-893-29 | Sequence 29, Appl |
| 34 | 131.5 | 12.2 | 399 | 16 | US-10-308-862-6 | Sequence 6, Appl |
| 35 | 128 | 11.9 | 484 | 10 | US-09-918-995-27921 | Sequence 27921, A |
| 36 | 93.5 | 8.7 | 1860 | 16 | US-10-369-493-41803 | Sequence 41803, A |
| 37 | 88 | 8.2 | 1248 | 16 | US-10-260-238-4369 | Sequence 4369, Ap |
| 38 | 87.5 | 8.1 | 2836 | 16 | US-10-080-334-59 | Sequence 59, Appl |
| 39 | 86.5 | 8.0 | 1157 | 13 | US-10-424-599-96351 | Sequence 96351, A |
| 40 | 86.5 | 8.0 | 7158 | 15 | US-10-156-761-201 | Sequence 201, App |
| 41 | 86.5 | 8.0 | 9025608 | 15 | US-10-156-761-1 | Sequence 1, Appl |
| 42 | 84.5 | 7.8 | 20986 | 13 | US-10-158-844-54 | Sequence 54, Appl |
| 43 | 84 | 7.8 | 2419 | 9 | US-09-893-238-8 | Sequence 8, Appl |
| 44 | 84 | 7.8 | 2836 | 16 | US-10-320-797-1301 | Sequence 1301, Ap |
| 45 | 84 | 7.8 | 4836 | 16 | US-10-320-797-301 | Sequence 301, App |

ALIGNMENTS

RESULT 1

US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

us-10-029-020-14_copy_2400_2600.p2n.rnpb

Mon Aug 16 09:01:13 2004

```

; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)...(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 3.33e-143 Length: 8354
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-383-201-43 (1-8354)
QY 1 IleGlyTyrHisGlyClyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
DB 7232 ATAGGCTACCATGGTGGCTCTATGATCCACTCACCAGCTGTCCACATGGCGCGGA 7291
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
DB 7292 GATTATGATGTGCTGGCGGAGCTGGACTAGCCAGACCCAGAGCTGTGGAAGCACCTT 7351
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
DB 7352 AGTAGCAGCAACGTATGCTTTTAACTCTATATGTTTCAAAAACCAACACCCATCAGC 7411
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
DB 7412 AACTCCAGGACATCAAGTGTTCATGACAGATGTTTAAACAGCTGGCTGTCACTTTGGA 7471
QY 81 PheGlnLeuHisAsnValMetProGlyTyrProLysProAspMetAspAlaMetGluPro 100
DB 7472 TTCCAGCTACACAACTGATCCCTGGTTATCCCAACCCAGACATGATGATCCATGGAAACC 7531
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
DB 7532 TCCTACGAGCTCATCCACACACAGATGAAACCCAGAGAGTGGGACACACAGTCTATC 7591
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
DB 7592 CTCGGGGTACAGTGTGAAGTACAGAGCAGCTCAAGGCCTTTGTTCACCTTAGAACGGTTT 7651
; RESULT 2

```

```

; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)...(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 3.33e-143 Length: 8354
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-383-201-43 (1-8354)
QY 1 IleGlyTyrHisGlyClyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
DB 7232 ATAGGCTACCATGGTGGCTCTATGATCCACTCACCAGCTGTCCACATGGCGCGGA 7291
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
DB 7292 GATTATGATGTGCTGGCGGAGCTGGACTAGCCAGACCCAGAGCTGTGGAAGCACCTT 7351
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
DB 7352 AGTAGCAGCAACGTATGCTTTTAACTCTATATGTTTCAAAAACCAACACCCATCAGC 7411
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
DB 7412 AACTCCAGGACATCAAGTGTTCATGACAGATGTTTAAACAGCTGGCTGTCACTTTGGA 7471
QY 81 PheGlnLeuHisAsnValMetProGlyTyrProLysProAspMetAspAlaMetGluPro 100
DB 7472 TTCCAGCTACACAACTGATCCCTGGTTATCCCAACCCAGACATGATGATCCATGGAAACC 7531
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
DB 7532 TCCTACGAGCTCATCCACACACAGATGAAACCCAGAGAGTGGGACACACAGTCTATC 7591
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
DB 7592 CTCGGGGTACAGTGTGAAGTACAGAGCAGCTCAAGGCCTTTGTCACTTAGAACGGTTT 7651
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
DB 7652 GACCAAGCTCATGGCTCCCAATCACCAGCTGCCAGAGCTCCAAAGACCAAGAGAGTTT 7711
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
DB 7712 GCATCCAGCGGCTCAGCTTTTGGCAAGGGGGTCAAGTTTGCCTTGAAGGATGCCAGATG 7771
QY 181 ThrThrAspIleLeuSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn 200
DB 7772 ACCACAGACATCATCAGTGTGGCCAAATGAGAGTGGGGGAGGGTGTCTGCCATCTTGAAC 7831
QY 201 His 201
DB 7832 CAT 7834

```


;; TITLE OF INVENTION: Using the Same
;; FILE REFERENCE: 21402-537
;; CURRENT APPLICATION NUMBER: US/10/042,865
;; CURRENT FILING DATE: 2002-05-17
;; PRIOR APPLICATION NUMBER: 60/260,417
;; PRIOR FILING DATE: 2001-01-09
;; PRIOR APPLICATION NUMBER: 60/260,831
;; PRIOR FILING DATE: 2001-01-10
;; PRIOR APPLICATION NUMBER: 60/272,338
;; PRIOR FILING DATE: 2001-02-28
;; PRIOR APPLICATION NUMBER: 60/274,876
;; PRIOR FILING DATE: 2001-03-09
;; PRIOR APPLICATION NUMBER: 60/284,704
;; PRIOR FILING DATE: 2001-04-18
;; NUMBER OF SEQ ID NOS: 264
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 1
;; LENGTH: 8438
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-042-865-1

Alignment Scores:
Pred. No.: 3,56e-141 Length: 8438
Score: 1063.00 Matches: 201
Percent Similarity: 96.17% Conservative: 0
Best Local Similarity: 96.17% Mismatches: 0
Query Match: 98.70% Indels: 8
DB: 13 Gaps: 1

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-042-865-1 (1-8438)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
DB 7259 ATAGGCTACCATGGTGGCTCTATGATCCACTCACCAAGCTTGTCCACATGGGCGCGGA 7318
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
DB 7319 GATTATGATGTGTGGCGGAGCGTGGCTAGCCAGACACAGAGCTGTGGAGACCTT 7378
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProLeuSer 60
DB 7379 AGTAGCAGCAAGCTCATGCTTTTAACTCTATATGTTCAAAAACCAACACCCATCAGC 7438
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
DB 7439 AACTCCAGGACATCAAGTGTCTCATGACAGATGTTTAAACAGCTGGCTGCTCACTTTGA 7498
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
DB 7499 TTCCAGCTACACAACGTGATCCCTGGTTATCCCAAAACACAGACATGATGCCATGGACCC 7558
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLys----- 118
DB 7559 TCCTACAGCTCATCCACACACAGATGAAACGAGGAGTGGGACACACAGCAAGTAAT 7618
QY 119 -----SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLys 132
DB 7619 CCTGCACAAAGGCTGCCAGCTCTATCCCTGGGGTACAGTGTGAATACAGACGAGCTCAAG 7678
QY 133 AlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGln 152
DB 7679 GCCTTTGTACCTTAGAACGGTTTGACAGCTCTATGGCTCCACATCACCAGCTGCCAG 7738
QY 153 GlnAlaProLysThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyValLys 172
DB 7739 CAGGCTCCAAAGACCAAGAGTGTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGGTCAAG 7798
QY 173 PheAlaLeuLysAspGlyArgValThrAspIleIleSerValAlaAsnGluAspGly 192
DB 7799 TTTGCCCTTGAAGGATGGCCGAGTGACACACAGACATCATCATGTGGCCCAATGAGGATGG 7858
QY 193 ArgArgValAlaAlaIleLeuAsnHis 201

Db 7859 CGAAGGTTGCTGCCATCTTGAACCAT 7885
RESULT 5
US-10-295-027-927
; Sequence 927, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 927
; LENGTH: 12880
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-927

Alignment Scores:
Pred. No.: 1.96e-78 Length: 12880
Score: 630.50 Matches: 116
Percent Similarity: 74.88% Conservative: 36
Best Local Similarity: 57.14% Mismatches: 46
Query Match: 58.54% Indels: 5
DB: 16 Gaps: 2

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-295-027-927 (1-12880)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
DB 7133 ATTTGTTTTCATGGAGGACTCTATGATTTCTTAAATTTAGTGCACCTGGGGCAAG 7192
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
DB 7193 GATTATGATGTTGTGTGGCAGATGGACCAACGCCCTATCATCATATGGAACAGTTG 7252
QY 41 SerSerSerAsnValMet-----ProPheAsnLeuTyrMetPheLysAsnAsnPro 58
DB 7253 -----AACCTCCTTCCTAAACCATCAACCTCTACTCTCTTTGAAATAATACCCCA 7303

QY 59 lIeSerAsnSerGlnAspIleLeuHisValAsnSerThrLeuThr 78
 Db 7304 GTTGGCAAAATTCAGAGTTCGCAAGTATACACAGACATCAGAGTGGTGGAGCTA 7363
 QY 79 PheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMet 98
 Db 7364 TTTGGTTTCCAAATACACAATGCTACTACCTGGATTCCCAAACTGAAATTAGAAAATTGA 7423
 QY 99 GluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLys 118
 Db 7424 GAATTAACCTTAGAGCTTCTACGGCTTCAGACAAACTCAAGAGTGGGATCCTGGAAAG 7483
 QY 119 SerLeuLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGlu 138
 Db 7484 ACTATCCTGGGCATTCAGTGTGAAGTCCAGAAACAGCTCAGGAATTCATTTCTTGAC 7543
 QY 139 ArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLys 158
 Db 7544 CAATCCTATGACTCCCGCATACATGATGAGCGGTGCTTGAAGGAGGGAAGCAACA 7603
 QY 159 LysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGly 178
 Db 7604 AGGTTTCTGCTGCTCCCTTCGTTTGGGAAAGGTATAAAATTTGCCATCAAGGATGGC 7663
 QY 179 ArgValThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAlaIle 198
 Db 7664 ATAGTAACAGCTGATATATAGGAGTAGCCCAATGAGATAGCAGCGCGCTTGCTGCCATT 7723
 QY 199 LeuAsnHis 201
 Db 7724 CTCAATAAT 7732

RESULT 6

US-10-342-887-1743
 ; Sequence 1743, Application US/10342887
 ; Publication No. US20040058340A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter S.
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Christopher J.
 ; APPLICANT: Van 't Veer, Laura Johanna
 ; APPLICANT: Van de Vijver, Marc J.
 ; APPLICANT: Bernards, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; CURRENT APPLICATION NUMBER: US/10/342,887
 ; CURRENT FILING DATE: 2003-01-15
 ; PRIOR APPLICATION NUMBER: 60/298,918
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 60/380,710
 ; PRIOR FILING DATE: 2002-05-14
 ; PRIOR APPLICATION NUMBER: 10/172,118
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 1743
 ; LENGTH: 3614
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-342-887-1743

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 5,07e-70 | Length: | 3614 |
| Score: | 566.00 | Matches: | 107 |
| Percent Similarity: | 71.08% | Conservative: | 38 |
| Best Local Similarity: | 52.45% | Mismatches: | 47 |
| Query Match: | 52.55% | Indels: | 12 |
| DB: | 13 | Gaps: | 4 |

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-342-887-1743 (1-3614)

QY 1 lIeGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
 Db 2192 ATTGGATTTCATGTGTGGCTGTATGACCACTACCAAAATTAATCCATTTTGGAGAAGA 2251
 QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
 Db 2252 GATTATGACATTTTGGCAGGACGGTGGACAACACCTGACATAGAAATCTGGAAGAAGATT 2311
 QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
 Db 2312 ---GGGAAGGAGCCAGCTCCCTTTTAACTTGTACATGTTTAGAATAACACCTCCAGC 2368
 QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrPheLeuThrPheGly 80
 Db 2369 AAAATCCATGACGTGAAAGATTACATCACAGATGTTAACAGCTGGCTGGTGCATTTGTT 2428
 QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 Db 2429 TTCCATCTGCACAATGCTATTTCCCTGGATTCCCTGTTCCCAAAATTTGATTTAACAGAACCT 2488
 QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
 Db 2489 TCTTACGAACCTTGTG-----AAGAGTCAGCAGTGGGATGATATACCCCCCATC 2536
 QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 Db 2537 TTCGAGTCAGCAGCAGCAAGTGGCGGCGAGCCCAAGGCCTTCTCTGCTGGGGAAGATG 2596
 QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
 Db 2597 GCCGAGGTGCAG-----GTGAGCGCGCGCGCGCGCGCGCGCGCTCTCGG 2644
 QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
 Db 2645 CTGTGGTTTCGCCACGCTCAAGTCGTCATCGGCAAGGGCTCATGCTGGCGCTCAGCCAG 2704
 QY 178 GlyArgValThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAla 197
 Db 2705 GCGCGCTGCAGACCAACGTCCTCAACATCGCCACGAGGACTGCATCAAGTGGCGGCC 2764
 QY 198 lIeLeuAsnHis 201
 Db 2765 GTGCTCAACAAC 2776

RESULT 7
 US-10-172-118-1743
 ; Sequence 1743, Application US/10172118
 ; Publication No. US20030224374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Chris
 ; APPLICANT: Van 't Veer, Laura
 ; APPLICANT: Van de Vijver, Marc
 ; APPLICANT: Bernards, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-175-999
 ; CURRENT APPLICATION NUMBER: US/10/172,118
 ; CURRENT FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 60/380,770
 ; PRIOR FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 1743
 ; LENGTH: 3614
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: NM_018104
 ; DATABASE ENTRY DATE: 2001-06-18
 US-10-172-118-1743

```

Alignment Scores:
Pred. No.: 5,07e-70 Length: 3614
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 13 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-172-118-1743 (1-3614)

Qy 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
Db 2192 ATTGGATTTCATGGTGGCTGATGACCACTACCAATTAATCCATTGGAGAAAGA 2251
Qy 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 2252 GATTATGACATTTGGCAGGCGGTGGACACACCTGCATAGAAATCTGGAAAGAATT 2311
Qy 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 2312 ---GGGAAGGCCAGCTCCCTTTAACTGTACATGTTTAGGAATAACACCTGCAAGC 2368
Qy 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuThrPheGly 80
Db 2369 AAAATCCATGACGTGAACATTAACATACAGATGTTAAAGCTGCTGTGACATTTGTT 2428
Qy 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 2429 TTCCATCTGCACATGCTATTCTGATTCCTGATTCCTGTTTCCAAATTTGATTTAAACAGAACCT 2488
Qy 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
Db 2489 TCTTACGAACTTGTG-----AAGAGTCAGCAGTGGGATGATATACGCCCATC 2536
Qy 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 2537 TTCGAGTTCAGCAGCAAGTGGCGGCGAGGCAAGGCTCTCTGTGCTGGGGAAGATG 2596
Qy 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 2597 GCGAGGTGCAG-----GTGAGCGCGCGCGCGCGCGCGCGCGCGCTCTGG 2644
Qy 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 2645 CTGTGTTTCGCCACGCTCAAGTCGCTGATCGGCAAGGCGCTCATGCTGCGCGTCACCCAG 2704
Qy 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAla 197
Db 2705 GCGCGGTGTCAGACCAACGCTGCTCAACATCGCCACGAGGACTGCATCAAGGTGGCGGCC 2764
Qy 198 IleLeuAsnHis 201
Db 2765 GTGCTCAACAAC 2776

RESULT 8
US-10-038-854-39
Sequence 39, Application US/10038854
Publication No. US20040022781A1
GENERAL INFORMATION:
APPLICANT: Li, Li
APPLICANT: Kimberly A
APPLICANT: Wolenc, Adam R
APPLICANT: Vernet, Corine
APPLICANT: Eissen, Andrew J
APPLICANT: Liu, Xiaohong
APPLICANT: Malyankar, Uriel M
APPLICANT: Shimkets, Richard A
APPLICANT: Tchernev, Velizar
APPLICANT: Spaderna, Steven K
APPLICANT: Gorman, Linda
APPLICANT: Kekuda, Ramesh
APPLICANT: Patturajan, Meera
APPLICANT: Gusev, Vladimiro Y

APPLICANT: Gangolli, Esha A
APPLICANT: Guo, Xiaofia S
APPLICANT: Shenoy, Suresh G
APPLICANT: Rastelli, Luca
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc
APPLICANT: Burgess, Catherine E
APPLICANT: Edinger, Shlomit R
APPLICANT: Ellerman, Karen
APPLICANT: Gunther, Erik
APPLICANT: Smithson, Glenda
APPLICANT: Millet, Isabelle
APPLICANT: MacDougall, John R
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-230
CURRENT APPLICATION NUMBER: US/10/038,854
CURRENT FILING DATE: 2003-01-22
PRIOR APPLICATION NUMBER: 60/258,928
PRIOR FILING DATE: 2000-12-29
PRIOR APPLICATION NUMBER: 60/259,415
PRIOR FILING DATE: 2001-01-02
PRIOR APPLICATION NUMBER: 60/259,785
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: 60/269,814
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/279,832
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,833
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,863
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/283,889
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,447
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/286,683
PRIOR FILING DATE: 2001-04-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 411
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 39
LENGTH: 8473
TYPE: DNA
ORGANISM: Homo sapiens
US-10-038-854-39

Alignment Scores:
Pred. No.: 2,03e-69 Length: 8473
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 17 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-038-854-39 (1-8473)

Qy 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
Db 7050 ATTGGATTTCATGGTGGCTGATGACCACTACCAATTAATCCATTGGAGAAAGA 7109
Qy 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 7110 GATTATGACATTTGGCAGGCGGTGGACACACCTGCATAGAAATCTGGAAAGAATT 7169
Qy 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7170 ---GGGAGGACCCAGCTCTCTTTAACTGTACATGTTTAGGAATAACACCTGCAAGC 7226
Qy 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuThrPheGly 80
Db 7227 AAAATCCATGACGTGAAAGATTATCATCACAGATGTTAAACAGCTGCTGCTGACATTTGTT 7286
Qy 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100

```

```
Db 7287 TTCCATCTGCACATGCTATTCTCGGATTCCCTGTTCCCAATTTGATTAAACAGAACT 7346
Qy 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7347 TCTTACGAACTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 7394
Qy 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7395 TTCGGAGTCCAGCAGCAAGTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7454
Qy 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7455 GCCAGGTGTCAG-----GTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7502
Qy 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 7503 CTGTGGTTCGCGCAGCGTCAAGTCGCTGATCGCGAAGGCGCTCATGCTGGCGTCAGCCAG 7562
Qy 178 GlyArgValThrThrAspIleIleSerValAlaAlaAsnGluAspGlyArgValAlaAla 197
Db 7563 GGCGCGTCGAGACCAAGGCTCAACATCGCCCAACGAGGACTGCATCAAGGTGGCGGCC 7622
Qy 198 IleLeuAsnHis 201
Db 7623 GTGCTCAACAAC 7634
```

RESULT 9

```
US-10-038-854-41
; Sequence 41, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytex, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennada
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-41

Alignment Scores:
Pred. No.: 2,04e-69 Length: 8487
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 17 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-038-854-41 (1-8487)

Qy 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7046 ATTGGATTTTCATGGTGGCTGTATGACCCCACTACCAAAATAATCCACTTTGGAGAAAGA 7105
Qy 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 7106 GATTATGACATTTTGGCAGGACGCTGACACACCTGACATAGATAATCTGGAAAAGAAAT 7165
Qy 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7166 ---GGGAAGGAGCCAGCTCTTTTAACTTTGATCATGTTTAGGAATAACACCCCTGCAAGC 7222
Qy 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7223 AAATCCATGACGTGAAGAGATTACATCAGATGTTTAAACAGCTGGCTGGTGCATTTGGT 7282
Qy 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7283 TTCCATCTGCACAAATGCTATTCTCGATTCCTGTTCCTGTTCCCAAAATTTGATTAAACAGAACCT 7342
Qy 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7343 TCTTACGAACTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 7390
Qy 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7391 TTCGGAGTCCAGCAGCAAGTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7450
Qy 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7451 GCCAGGTGTCAG-----GTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7498
Qy 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 7499 CTGTGGTTCGCGCAGCGTCAAGTCGCTGATCGCAAGGCGCTCATGCTGGCGCTCAGCCAG 7558
Qy 178 GlyArgValThrThrAspIleIleSerValAlaAlaAsnGluAspGlyArgValAlaAla 197
Db 7559 GGCGCGTTCGAGACCAACAGTCTCAACATCGCCCAACGAGGACTGCATCAAGGTGGCGGCC 7618
Qy 198 IleLeuAsnHis 201
Db 7619 GTGCTCAACAAC 7630
```

| | | |
|--|---------|--|
| US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-038-854-37 (1-8645) | | |
| QY | 1 | IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20 |
| Db | 7222 | ATTGGATTTCATGGTGGCCTGTATGACCCACTACCAAAATTAATCCACTTTGGAGAAGA 7281 |
| QY | 21 | AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40 |
| Db | 7282 | GATTATGACATTTTGGCAGGAGCGTGGACAACACTGCATAGAAATCTCGAAAGAATT 7341 |
| QY | 41 | SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60 |
| Db | 7342 | ---GGGAAGGACCCAGCTCCTTTTAACTTTGTACATGTTTAGGAATAACAACCCCTGCAGC 7398 |
| QY | 61 | AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrProLeuThrPheGly 80 |
| Db | 7399 | AAATCCATGACGTGAAGAATTACATCACAGATGTTAACAGCTGGCTGGTGCACATTGGT 7458 |
| QY | 81 | PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100 |
| Db | 7459 | TTCCATCTGCACAATGCTATTCTCGATTCTCCTGTTCCCAAAATTTGATTTAACAGAACCT 7518 |
| QY | 101 | SerTyrGluLeuIleHisThrGlnMetLysThrGlnMetLysThrGlnMetLysThrGlnMetLys 120 |
| Db | 7519 | TCATTACGAACCTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 7566 |
| QY | 121 | LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140 |
| Db | 7567 | TTGGAGTCCAGCAGCAGTGGCGCGGCGAGCCCTTCTCTGCTGGGAAGATG 7626 |
| QY | 141 | AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159 |
| Db | 7627 | GCGAGGTGCAG-----GTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7674 |
| QY | 160 | -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177 |
| Db | 7675 | CTGTGGTTCCGCCACCGTCAAGTCGCTGATCGCAAGCGCGCTCATCTGCGCGCTCAGCCAG 7734 |
| QY | 178 | GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAla 197 |
| Db | 7735 | GGCCGCGTGCAGACCAACAGTGTCTCAACATCGCAACGAGGACTGCATCAAGGTGGCGGC 7794 |
| QY | 198 | IleLeuAsnHis 201 |
| Db | 7795 | GTGCTCAACAC 7806 |
| RESULT 11 | | |
| US-10-038-854-35 | | |
| ; Sequence 35, Application US/10038854 | | |
| ; Publication NO. US20040022781A1 | | |
| ; GENERAL INFORMATION: | | |
| ; APPLICANT: Spytek, Kimberly A | | |
| ; APPLICANT: Li, Li | | |
| ; APPLICANT: Wolenc, Adam R | | |
| ; APPLICANT: Vernet, Corine | | |
| ; APPLICANT: Eisen, Andrew J | | |
| ; APPLICANT: Liu, Xiaohong | | |
| ; APPLICANT: Malyankar, Uriel M | | |
| ; APPLICANT: Shinkets, Richard A | | |
| ; APPLICANT: Tchernev, Velizar | | |
| ; APPLICANT: Spaderna, Steven K | | |
| ; APPLICANT: Gorman, Linda | | |
| ; APPLICANT: Kekuda, Ramesh | | |
| ; APPLICANT: Patturajan, Meera | | |
| ; APPLICANT: Gangolli, Esha A | | |
| ; APPLICANT: Guo, Xiaojia S | | |
| ; APPLICANT: Shenoy, Suresh G | | |
| ; APPLICANT: Rastelli, Luca | | |
| ; APPLICANT: Casman, Stacie J | | |
| ; APPLICANT: Boldog, Ferenc | | |
| ; APPLICANT: Burgess, Catherine E | | |
| ; APPLICANT: Edinger, Shlomit R | | |
| ; APPLICANT: Ellerman, Karen | | |
| ; APPLICANT: Gunther, Erik | | |
| ; APPLICANT: Smithson, Glennnda | | |
| ; APPLICANT: Millet, Isabelle | | |
| ; APPLICANT: Macdougall, John R | | |
| ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same | | |
| ; FILE REFERENCE: 21402-230 | | |
| ; CURRENT APPLICATION NUMBER: US/10/038,854 | | |
| ; CURRENT FILING DATE: 2003-01-22 | | |
| ; PRIOR APPLICATION NUMBER: 60/258,928 | | |
| ; PRIOR FILING DATE: 2000-12-29 | | |
| ; PRIOR APPLICATION NUMBER: 60/259,415 | | |
| ; PRIOR FILING DATE: 2001-01-02 | | |
| ; PRIOR APPLICATION NUMBER: 60/259,785 | | |
| ; PRIOR FILING DATE: 2001-01-04 | | |
| ; PRIOR APPLICATION NUMBER: 60/269,814 | | |
| ; PRIOR FILING DATE: 2001-02-20 | | |
| ; PRIOR APPLICATION NUMBER: 60/279,832 | | |
| ; PRIOR FILING DATE: 2001-03-29 | | |
| ; PRIOR APPLICATION NUMBER: 60/279,833 | | |
| ; PRIOR FILING DATE: 2001-03-29 | | |
| ; PRIOR APPLICATION NUMBER: 60/279,863 | | |
| ; PRIOR FILING DATE: 2001-03-29 | | |
| ; PRIOR APPLICATION NUMBER: 60/283,889 | | |
| ; PRIOR FILING DATE: 2001-04-13 | | |
| ; PRIOR APPLICATION NUMBER: 60/284,447 | | |
| ; PRIOR FILING DATE: 2001-04-18 | | |
| ; PRIOR APPLICATION NUMBER: 60/286,683 | | |
| ; PRIOR FILING DATE: 2001-04-25 | | |
| ; Remaining Prior Application data removed - See File Wrapper or PALM. | | |
| ; NUMBER OF SEQ ID NOS: 411 | | |
| ; SOFTWARE: PatentIn Ver. 2.1 | | |
| ; SEQ ID NO 37 | | |
| ; LENGTH: 8645 | | |
| ; TYPE: DNA | | |
| ; ORGANISM: Homo sapiens | | |
| US-10-038-854-37 | | |
| Alignment Scores: | | |
| Pred. No.: | 2,1e-69 | 8645 |
| Score: | 566.00 | 107 |
| Percent Similarity: | 71.08% | 38 |
| Best Local Similarity: | 52.45% | 47 |
| Query Match: | 52.55% | 12 |
| DB: | 17 | 4 |

```
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35

Alignment Scores:
Pred. No.: 2,11e-69 Length: 8675
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 17 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-038-854-35 (1-8675)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
Db 7234 ATTGGATTTTCATGGTGGCGCTGTATGACCCCACTACCAAAATTAATCCACTTTGGAGAAAGA 7293

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuLeuTyrLysHisLeu 40
Db 7294 GATTATGACATTTTGGCAGGCGGTGGACAACTGACATAGAAATCTGGAAAGAAAT 7353

QY 41 SerSerSerValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7354 ---GGGAAGGACCCAGCTCTTTTAACTTGATGATGTTTAGGAATAACACCCCTGCAAGC 7410

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7411 AAAATCCATGAGCTGAAGATTTACATACAGATGTTTAAACAGCTGCTGTGACATTTGGT 7470

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7471 TTCCATCTGCACAAATGCTATTCCTGGATTCCCTGTTCCCAATTTGATTTAACAGAACT 7530

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerIleSer 120
Db 7531 TCTTACGAACTTGCTG ---AGAGTCAGCAGTGGGATGATATACCGCCCATC 7578

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7579 TTCGGAGTCCAGCAGCAAGTGGCGCGGCAAGGCTTCTCTGCTGCTGGGAGATG 7638

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7639 GCCGAGGTGCAG-----GTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCTGG 7686

QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 7687 CTGTGGTTGCGCCACGCTCAAGTCTGATCGCGCAGGCGCTCATCTGCTGGCGTCAGCCAG 7746

QY 178 GlyArgValThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAla 197
Db 7747 GGCGCGTGCAGACCAACAGTCTCAACATCGCCACGAGGACTGATCAAGGTGGCGGCC 7806

QY 198 IleLeuAsnHis 201
Db 7807 GTGCTCAACAAC 7818

RESULT 12
US-09-823-245A-85
; Sequence 85, Application US/09823245A
; Publication No. US20020039760A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fichtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakari
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6401
; CURRENT APPLICATION NUMBER: US/09/823,245A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/194,941
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 791
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 566,604,734,779
; OTHER INFORMATION: n = a,t,g, or c
US-09-823-245A-85

Alignment Scores:
Pred. No.: 1,18e-69 Length: 791
Score: 556.00 Matches: 106
Percent Similarity: 71.14% Conservative: 37
Best Local Similarity: 52.74% Mismatches: 52
Query Match: 51.62% Indels: 7
DB: 13 Gaps: 3

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-823-245A-85 (1-791)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
Db 102 ATTGGATTTTCATGGTGGCGCTGTATGACCCCACTACCAAAATTAATCCACTTTGGAGAAAGA 161

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 162 GATTATGACATTTTGGCAGGCGGTGGACAACTGACATAGAAATCTGGAAAGAAAT 221

QY 41 SerSerSerValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 222 ---GGGAAGGACCCAGCTCTTTTAACTTGATGATGTTTAGGAATAACACCCCTGCAAGC 278
```

| | | |
|------|--|------|
| 7307 | NTTGGCTTCCAGCGAGGCGCTCTATGATGACCCCTCACCAAGCTCGTCCACTTTTACTCAACGT | 7366 |
| 21 | AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuLeuTrpLysHisLeu | 40 |
| 7367 | GATTATGACGTCTGGCAGAGCGGTGGAGCTCCCCGACTACCACTGTGGAGGAACGTG | 7426 |
| 41 | SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProLysSer | 60 |
| 7427 | GGCAAGGACCGACGCC---CCCTTCAACCTGTACA GTTCAAGAACAAACAATCCTCTGAGC | 7483 |
| 61 | AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly | 80 |
| 7484 | AATGACCTGGACTTAAAGAACTACGTGCAGACAGCTGAAGAGCTGGCTTGTGATGTTTGA | 7543 |
| 81 | PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro | 100 |
| 7544 | TTTCAGCTCAGCAACAATCATCTCTGTGATCTCCGAGAGCCAAATGATTGTTGGCTCC | 7603 |
| 101 | SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle | 120 |
| 7604 | CCCTATGAACCTGTCAGAGAGTCAAGCAAGC-----GAGAACGCGACAGCTCAT | 7651 |
| 121 | LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluAArgPhe | 140 |
| 7652 | ACAGGTGTCAGCAGACAACTGAGAGCGCATAAACAGCGCCTCTCTCGGAA----- | 7705 |
| 141 | AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- | 159 |
| 7706 | -----GGACAGGTGCATCACTATAAAGCTCCATCCAGCATCCGAGAGAAAGCA | 7753 |
| 160 | -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLys | 176 |
| 7754 | GGCCATGGTTTGTCTACCACCAACCATCATCGGCAAGGCATCATGTTTGCCATCAA | 7813 |
| 177 | AspGlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAla | 196 |
| 7814 | GAAGGGCGGGTGACCACAGGAGTGTCTAGCATCCAGTGAAGCAGCCCGCAAGGTAGCA | 7873 |
| 197 | AlaIleLeuAsnHis | 201 |
| 7874 | TCCGTGTTGAACAT | 7888 |

RESULT 14
 US-09-808-602-77
 ; Sequence 77, Application US/09808602
 ; Patent No. US20020155115A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vernet, Corine A
 ; APPLICANT: Fernandes, Elma
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Herrman, John L
 ; APPLICANT: Majumder, Kumud
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Mezes, Peter S
 ; APPLICANT: MacDougall, John
 ; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding
 ; FILE REFERENCE: 15966-697 CIP
 ; CURRENT APPLICATION NUMBER: US/09/808,602
 ; CURRENT FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 09/800,198
 ; PRIOR FILING DATE: 2001-03-05
 ; PRIOR APPLICATION NUMBER: 60/186,596
 ; PRIOR FILING DATE: 2000-03-03
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 77
 ; LENGTH: 8797
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 US-09-808-602-77

```

; PRIOR APPLICATION NUMBER: 807188,350
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-808-602-77

Alignment Scores: 3.93e-65 Length: 8797
Pred. No.:

```


Score: 536.50 Matches: 104
Percent Similarity: 68.78% Conservative: 37
Best Local Similarity: 50.73% Mismatches: 49
Query Match: 49.81% Indels: 15
DB: 9 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-808-602-77 (1-8797)

QY 1 IleglyTyHisGlyGlyLeuTyAspProLeuThyHisValHisMetGlyArgArg 20
Db 7307 ATTGGCTTCCAGGAGGCTCTATGACCCCTTACCAAGCTGCTCCACTTTACTCAACGT 7366

QY 21 AspTyAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7367 GATTATGACGTGCTGGCAGGAGCGTGGACGCTCCCGGACTACACCATGTGGAGAACGTG 7426

QY 41 SerSerSerAsnValMetProPheAsnLeuTyMetPheLysAsnAsnProIleSer 60
Db 7427 GGCAGGAGCCAGCC---CCCTTCAACCTGTACATGTTCAAGAACAAATCTCTGAGC 7483

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThyPheGly 80
Db 7484 AATGAGCTGGACTTAAGAACTAGTGACAGCTGAAGCTGGCTGTGATGTTTGA 7543

QY 81 PheGlnLeuHisAsnValIleProGlyTyTrpProLysProAspMetAspAlaMetGluPro 100
Db 7544 TTTGAGCTCAGCAACATCTCTGATTCCTGGATTCCCGAGGCCAAATGATTTTGGCTCCC 7603

QY 101 SerTyGlnLeuLeuHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7604 CCTTATGAACCTGCAGAGCTCAAGCAAGC-----GAGAACGGCAGCTCATT 7651

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7652 ACAGGTGTCAGCAGCAACTGAGAGCATAAACGAGGCTTCTGGCTCTGGA 7705

QY 141 AspGlnLeuTyGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7706 -----GGACAGGTCTACCTATAAAGCTCCATGCGCAAGCAAGC-----GAGAACGGCAGCTCATT 7651

QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLys 176
Db 7754 GGCCACTGGTTTCTACCAACCATCCATCGCAAGGCTCCTGGCTCTGGA 7813

QY 177 AspGlyArgValThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAla 196
Db 7814 GAAGGGCGGTGACCAAGAGTGTCTAGCATCGCCAGTGGAGGACGCCCAAGGTAGCA 7873

QY 197 AlaIleLeuAsnHis 201
Db 7874 TCCGTGTTGAACAAT 7888

RESULT 15

US-09-800-198-62
; Sequence 62, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 62
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-62

Alignment Scores:

Pred. No.: 3,93e-65 Length: 8797
Score: 536.50 Matches: 104
Percent Similarity: 68.78% Conservative: 37
Best Local Similarity: 50.73% Mismatches: 49
Query Match: 49.81% Indels: 15
DB: 10 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-800-198-62 (1-8797)

QY 1 IleglyTyHisGlyGlyLeuTyAspProLeuThyHisValHisMetGlyArgArg 20
Db 7307 ATTGGCTTCCAGGAGGCTCTATGACCCCTTACCAAGCTGCTCCACTTTACTCAACGT 7366

QY 21 AspTyAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7367 GATTATGACGTGCTGGCAGGAGCGTGGACGCTCCCGGACTACACCATGTGGAGAACGTG 7426

QY 41 SerSerSerAsnValMetProPheAsnLeuTyMetPheLysAsnAsnProIleSer 60
Db 7427 GGCAGGAGCCAGCC---CCCTTCAACCTGTACATGTTCAAGAACAAATCTCTGAGC 7483

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThyPheGly 80
Db 7484 AATGAGCTGGACTTAAGAACTAGTGACAGCTGAAGCTGGCTGTGATGTTTGA 7543

QY 81 PheGlnLeuHisAsnValIleProGlyTyTrpProLysProAspMetAspAlaMetGluPro 100
Db 7544 TTTGAGCTCAGCAACATCTCTGATTCCTGGATTCCCGAGGCCAAATGATTTTGGCTCCC 7603

QY 101 SerTyGlnLeuLeuHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7604 CCTTATGAACCTGCAGAGCTCAAGCAAGC-----GAGAACGGCAGCTCATT 7651

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7652 ACAGGTGTCAGCAGCAACTGAGAGCATAAACGAGGCTTCTGGCTCTGGA 7705

QY 141 AspGlnLeuTyGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7706 -----GGACAGGTCTACCTATAAAGCTCCATGCGCAAGCAAGC-----GAGAACGGCAGCTCATT 7651

QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLys 176
Db 7754 GGCCACTGGTTTCTACCAACCATCCATCGCAAGGCTCCTGGCTCTGGA 7813

QY 177 AspGlyArgValThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAla 196
Db 7814 GAAGGGCGGTGACCAAGAGTGTCTAGCATCGCCAGTGGAGGACGCCCAAGGTAGCA 7873

QY 197 AlaIleLeuAsnHis 201
Db 7874 TCCGTGTTGAACAAT 7888

Search completed: August 14, 2004, 19:22:42
Job time : 457.845 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 59.4656 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_2400_2600
Perfect score: 1077
Sequence: 1 IGHHGLYDLPLTKLVHMR.....TDIISVANEDGRRVAALNH 201

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CNG 1.1-258 @runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| C 1 | 84.5 | 7.8 | 20386 | 4 | US-08-961-527-54 |
| 2 | 84 | 7.8 | 2419 | 3 | US-09-245-041-8 |
| 3 | 84 | 7.8 | 8827 | 3 | US-09-245-041-1 |
| 4 | 82.5 | 7.7 | 1650 | 4 | US-09-543-681A-548 |
| 5 | 81 | 7.5 | 45175 | 4 | US-09-453-702B-116 |
| C 6 | 81 | 7.5 | 49795 | 4 | US-09-453-702B-116 |
| 7 | 80 | 7.4 | 2280 | 4 | US-09-328-352-115 |
| C 8 | 80 | 7.4 | 9909 | 4 | US-08-961-527-12 |
| 9 | 78.5 | 7.3 | 3215 | 4 | US-09-620-312D-86 |
| 10 | 78 | 7.2 | 3254 | 1 | US-08-162-809-15 |
| 11 | 77.5 | 7.2 | 5706 | 4 | US-09-738-946-11 |
| 12 | 76.5 | 7.1 | 1119 | 4 | US-09-252-991A-2392 |
| | | | | | Sequence 54, Appli |
| | | | | | Sequence 8, Appli |
| | | | | | Sequence 1, Appli |
| | | | | | Sequence 548, App |
| | | | | | Sequence 116, App |
| | | | | | Sequence 60, Appl |
| | | | | | Sequence 115, App |
| | | | | | Sequence 12, Appl |
| | | | | | Sequence 86, Appl |
| | | | | | Sequence 15, Appl |
| | | | | | Sequence 11, Appl |
| | | | | | Sequence 2392, Ap |

| | | | | | | |
|------|------|-----|---------|---|----------------------|--------------------|
| 13 | 76.5 | 7.1 | 2772 | 4 | US-09-717-926-3 | Sequence 3, Appli |
| 14 | 76.5 | 7.1 | 7220 | 4 | US-09-717-926-1 | Sequence 1, Appli |
| 15 | 76 | 7.1 | 2745 | 4 | US-09-817-514A-5 | Sequence 5, Appli |
| 16 | 75.5 | 7.0 | 3132 | 4 | US-08-851-567B-60 | Sequence 60, Appli |
| 17 | 75 | 7.0 | 41708 | 4 | US-09-470-512A-3 | Sequence 3, Appli |
| C 18 | 74.5 | 6.9 | 1044 | 3 | US-09-202-832-2 | Sequence 2, Appli |
| C 19 | 74.5 | 6.9 | 1044 | 3 | US-09-202-832-11 | Sequence 11, Appli |
| C 20 | 74.5 | 6.9 | 1247 | 3 | US-09-202-832-3 | Sequence 3, Appli |
| 21 | 74.5 | 6.9 | 1533 | 4 | US-09-134-000C-3230 | Sequence 3230, Ap |
| C 22 | 74.5 | 6.9 | 1873 | 4 | US-09-016-434-1437 | Sequence 1437, Ap |
| 23 | 74.5 | 6.9 | 2397 | 4 | US-09-134-001C-2080 | Sequence 2080, Ap |
| 24 | 74.5 | 6.9 | 3852 | 4 | US-09-245-248B-29 | Sequence 29, Appl |
| C 25 | 74 | 6.9 | 11823 | 4 | US-08-956-171E-136 | Sequence 136, App |
| C 26 | 73.5 | 6.8 | 1309 | 1 | US-07-926-788A-1 | Sequence 1, Appli |
| 27 | 73.5 | 6.8 | 2625 | 3 | US-09-545-041-18 | Sequence 18, Appl |
| 28 | 73.5 | 6.8 | 2895 | 4 | US-09-543-681A-863 | Sequence 863, App |
| 29 | 73.5 | 6.8 | 4072 | 3 | US-09-245-041-16 | Sequence 16, Appl |
| 30 | 73.5 | 6.8 | 8589 | 3 | US-09-245-041-14 | Sequence 14, Appl |
| 31 | 73 | 6.8 | 720 | 4 | US-09-252-991A-13392 | Sequence 13392, A |
| 32 | 73 | 6.8 | 1296 | 4 | US-09-527-058-5 | Sequence 5, Appli |
| 33 | 73 | 6.8 | 3540 | 4 | US-09-107-532A-2472 | Sequence 2472, Ap |
| 34 | 73 | 6.8 | 1664976 | 4 | US-08-916-421B-1 | Sequence 1, Appli |
| 35 | 72.5 | 6.7 | 849 | 4 | US-09-134-000C-570 | Sequence 570, App |
| 36 | 72.5 | 6.7 | 4849 | 2 | US-08-540-804-13 | Sequence 13, Appl |
| 37 | 72.5 | 6.7 | 4849 | 2 | US-08-218-265-13 | Sequence 13, Appl |
| 38 | 72.5 | 6.7 | 4849 | 3 | US-08-521-872-13 | Sequence 13, Appl |
| 39 | 72.5 | 6.7 | 4849 | 3 | US-08-590-399-13 | Sequence 13, Appl |
| 40 | 72.5 | 6.7 | 37948 | 3 | US-09-251-645-11 | Sequence 11, Appl |
| C 41 | 72.5 | 6.7 | 118067 | 4 | US-09-497-855A-32 | Sequence 32, Appl |
| 42 | 72 | 6.7 | 1347 | 4 | US-09-134-001C-1309 | Sequence 1309, App |
| 43 | 72 | 6.7 | 2202 | 4 | US-08-956-171B-284 | Sequence 284, App |
| C 44 | 72 | 6.7 | 2329 | 1 | US-08-253-785-2 | Sequence 2, Appli |
| 45 | 72 | 6.7 | 2920 | 4 | US-08-976-259-10 | Sequence 10, Appl |

ALIGNMENTS

RESULT 1

US-08-961-527-54/c

; Sequence 54, Application US/08961527

; Patent No. 6420135

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,527

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 54:

; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 20986 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-54

Alignment Scores:
Pred. No.: 24.1 Length: 20986
Score: 84.50 Matches: 31
Percent Similarity: 43.01% Conservative: 9
Best Local Similarity: 33.33% Mismatches: 38
Query Match: 7.85% Indels: 15
DB: 4 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-08-961-527-54 (1-20986)

QY 34 HisGluLeuTrpLysHisLeuSerSerAsnValMetProPheAsnLeuTyrMetPhe 53
DB 13300 CATACAGAGAAAGCATCTACTATCCAA-----TTC 13268

QY 54 LysAsn-----AsnAsnProIleSerAsnSerGlnAsnLysCysPheMetThrAsp 71
DB 13267 AAGATATCTTAGGAATGACATTCGAATCTCTTCT-----TTTTAACAGCA 13217

QY 72 ValAsnSerTrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIleProGlyTyrPro 91
DB 13216 CTGACTTCTTTAACAAAGAACGTTTGTGTTTACAAATCTAGTTCCTTATTAAT 13157

QY 92 LysProAspMetAspAlaMetGluProSerTyrGluLeuIleHis---ThrGlnMetLys 110
DB 13156 TATTACGAACATTATGGGATTGAAGCGACAGCCACTGTATCATTTTCGAAGCAATT 13097

QY 111 ThrGlnGluTrpAspAsnSerLysSerIleLeuGlyVal 123
DB 13096 GTTCAAGAGTGGATGACTGAGCAGAGAAGATTGAAGGAGTT 13058

RESULT 2
US-09-245-041-8
; Sequence 8, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2419
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-8

Alignment Scores:
Pred. No.: 0.956 Length: 2419
Score: 84.00 Matches: 54
Percent Similarity: 37.67% Conservative: 30
Best Local Similarity: 24.22% Mismatches: 86
Query Match: 7.80% Indels: 53
DB: 10 Gaps: 10

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-245-041-8 (1-2419)

QY 2 GlyTyr---HisGlyLeuTyrAspProLeuThrLys-----LeuValHisMetGly 18
DB 1764 GGTATGGCCACACAGTAGTGTATTATGATGACAGGACCAAGGCTCTGTACGTTTCAGTGGC 1823

QY 19 ArgArg-----AspTyrAspVal 24
DB 1824 TACAAGGCTTTCAGCGCCAAATACCGCTTTCAGATGACCTCTACAGATACGATGTG 1883

QY 25 LeuAlaGlyArgTrpThr---SerProAspHisGluLeuTrpLysHisLeuSerSer 43
DB 1884 GATACTCAGATGTGGACCATTTCTTAAGCAGACGCGATTTTTCGTTTACTTCATACAGCT 1943

QY 44 AsnValMetProPheAsnLeuTyrMetPheLysAsnAsn-----AsnProIleSerAsn 61
DB 1944 GTGATAGTGTAGTGGAACCATCTGCTGTGTTTTCGAGGGAACACACACATGACACTTCCATG 2003

QY 62 SerGlnAspIleLysCysPheMetThrAsp-----ValAsnSer 74
DB 2004 AGCCACGGTCCCAATGCTTCTCTCGGACTTCATGCTTATGACATTTGCTTGTGACCGA 2063

QY 75 TrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAsp 94
DB 2064 TGGTCAGTGTCTCCAGACCTGAGCTCCATCAT-----GAT 2099

QY 95 MetAspAlaMetGluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrp 114
DB 2100 GTCACAGATTTGGCCATTGACGAGTCTTGTACACACAGCACCATGATGTGTCGGCGC 2159

QY 115 AspAsnSerLysSerIleLeuGlyVal-----GlnCysGluValGln 128
DB 2160 TTCACAGCTCTCTCTCAGTGACGCTCTTGTCTTTACCTCGGAGCAGTGGGATGCACAC 2219

QY 129 LysGlnLeuLysAlaPheValThr-----LeuGluArgPheAspGlnLeuTyr 144
DB 2220 CGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGACACACAG 2279

QY 145 GlySerThrIleThrSerCysGlnGlnAlaPro-LysThrLysLysPheAlaSerSerG1 164
DB 2280 TCGTCTCGATGATACCTCTCTGGAGTGGCAACTGAAGAACAGCAGAGAAAGTTAAATCA 2339

QY 164 ySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgValThrAspI1 184
DB 2340 GAGTGTGTTT-----CTAAAGAACCCCTTGACCATGACAGATGT 2378

QY 184 elleSer 186
DB 2379 GACCAGC 2385

RESULT 3
US-09-245-041-1
; Sequence 1, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-1

Alignment Scores:
Pred. No.: 7.24 Length: 8827
Score: 84.00 Matches: 54
Percent Similarity: 37.67% Conservative: 30

```



```

; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER: 60/110,955
; APPLICATION DATE: 04-DEC-1998
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45175
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 116:
US-09-453-702B-116

Alignment Scores:
Pred. No.: 233 Length: 45175
Score: 81.00 Matches: 44
Percent Similarity: 39.25% Conservative: 29
Best Local Similarity: 23.66% Mismatches: 75
Query Match: 7.52% Indels: 38
DB: 4 Gaps: 9

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-453-702B-116 (1-45175)

QY 23 AspValLeuAlaGly-----ArgTrpThrSer----- 31
Db 42089 GATGGGATCGAGAAACATTGACGGTACCATCGCTGGCGCAAAATGAGGTTATC 42148
QY 32 ---ProAspHisGluLeuTrpLysHisLeuSerSerSer-----AsnVal 45
Db 42149 GACAAGGATCGGAGAGCTGGGAAGCGTATCTCAGCATCATCGGCAAAACAGCCGCAAGCA 42208
QY 46 MetProPheAsn-----LeuTyrMetPheLysAsnAsnAsnProLysSerAsnSer 62
Db 42209 CTGGCTATGACCGTCAGACTATTTGGCCCTGTTGGCGAAGCTCGGATCATATTCAC 42268
QY 63 GlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGlyPheGln 82
Db 42269 AAAGATCCGCTGCACTGCAACAAATATATCATGTAATATCTGACGCAAAAGGCGGTGTTT 42328
QY 83 LeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluProSerTyr 102
Db 42329 GAACATGAAGAAACA-----GACCAGAGCTCTACTGATGCTCTCCAGCGTCAGCA 42379
QY 103 GluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIleLeuGly 122
Db 42380 GCACAACT--GCTCCAGTGGAGACGCGCAAGATCCGATATCTGACGCAAAAGGCGGTGTTT 42436
QY 123 ValGlnCysGluValGlnLysGlnLeuLysAla-----PheValThrLeuGluArg 139
Db 42437 GTGGAAGCTGAACCACTGTAGAGCGGTGAAGGACCAATTTATTCGCTTTACCGATAAG 42496
QY 140 PheAspGlnLeuTrpGlySerThrIleThrSerCysGlnGlnAlaProLys----- 156
Db 42497 GCGGGGAAATACGCG-----AGGCCAAACAACTTTCTGGT 42535
QY 157 ---ThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeu 175
Db 42536 CTGGACAAAGCGCTGGCTGCGCGGTACCGAAATCTCAAAGAGAATATTTTGCCCGA 42595
QY 176 LysAspGlyArgValThr 181
Db 42596 AAAAATGGCACATACAG 42613

; US-09-453-702B-60/c
; Sequence 60, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49795
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-453-702B-60

Alignment Scores:
Pred. No.: 271 Length: 49795
Score: 81.00 Matches: 44
Percent Similarity: 39.25% Conservative: 29
Best Local Similarity: 23.66% Mismatches: 75
Query Match: 7.52% Indels: 38
DB: 4 Gaps: 9

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-453-702B-60 (1-49795)

QY 23 AspValLeuAlaGly-----ArgTrpThrSer----- 31
Db 3137 GATGGATGCGAGAAACATTGACGGTACCATCGCTGGCGCAAAATGAGGTTATC 3078
QY 32 ---ProAspHisGluLeuTrpLysHisLeuSerSerSer-----AsnVal 45
Db 3077 GACAAGGATCGGAGAGCTGGGAAGCGTATCTCAGCATCAATGCGCAAAACAGCCGCAAGCA 3018
QY 46 MetProPheAsn-----LeuTyrMetPheLysAsnAsnAsnProLysSerAsnSer 62
Db 3017 CTGGCTATGACCGTCAGACTATTTGGCCCTGTTGGCGAAGCTCGGATCATATTCAC 2958
QY 63 GlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGlyPheGln 82
Db 2957 AAAGATCCGCTGCACTGCAACAAATATATCATGTAATATCTGACGCAAAAGGCGGTGTTT 2898

```

QY 83 LeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluProSerTyr 102
DB 2897 GAACATGAAGAAACA-----GACCAGAGCTCTACTGATGCTCTCCAGCGTCAGCA 2847
QY 103 GluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIleLeuGly 122
DB 2846 GCACAAACT---GCTCCAGTGGAGACGGCAGAAATCCGATCTCAAAAAAATGAATCCTG 2790
QY 123 ValGlnCysGluValGlnLysGlnLeuLysAla-----PheValThrLeuGluArg 139
DB 2789 GTGAGAGCTGAACCATCTGTAGAGCGTGAAGACCATTTTATTCGTCTTACCGATAAG 2730
QY 140 PheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLys----- 156
DB 2729 GCGGGGAAAAATACGGC-----AGGCAACAAACTTCTCGT 2691
QY 157 ---ThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeu 175
DB 2690 CTGACAGAGCGCTGGCTGCGCGGGTACCGAAATCTCAAAAGAAGAAATATTTGCCCGA 2631
QY 176 LysAspGlyArgValThr 181
DB 2630 AAAAATGSCACATACAG 2613
RESULT 7
US-09-328-352-115
; Sequence 115, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 115
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
; NAME/KEY: unsure
; LOCATION: (2238)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknd
US-09-328-352-115

Alignment Scores:
Pred. No.: 2.95 Length: 2280
Score: 80.00 Matches: 27
Percent Similarity: 43.33% Conservative: 12
Best Local Similarity: 30.00% Mismatches: 41
Query Match: 7.43% Indels: 10
DB: 4 Gaps: 3

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-328-352-115 (1-2280)

QY 116 AsnSerLysSerIleLeuGly-----ValGlnCysGluValGlnLysGln 130
DB 1564 AATAGCAAGCAGCAACAGGTAATGATTCATTCAGCTAGATGATCGAAAAAGTCT 1623
QY 131 LeuLysAlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSer 150
DB 1624 GTGAGCGCTGTACACAACTTGACCAATGGCTAAAGATACGGCTTTGTCATTCGGTT 1683
QY 151 CysGlnGlnAlaProLysThrLysLysPheAlaSerSerGly-----SerVal 166
DB 1684 CTGTAGTGTACAGAAGAGTCTGCTGCTGCAACCGGGGGCATATTCATGATCTGTG 1743
QY 167 PheGly---LysGlyValLysPheAlaLeuLysAspGlyArgValThrThrAspIlelle 185
DB 1744 CTGGCTACAAGGCACACATGCAATTAAGATGCAAAATGCAGAGCTGATATTGTT 1803
QY 186 SerValAlaAsnGluAspGlyArgVal 195

DB 1804 CAGAAAGCAAAATGATGAAGCTACAAAAATT 1833
RESULT 8
US-08-961-527-12/c
; Sequence 12, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9909 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-12

Alignment Scores:
Pred. No.: 29.4 Length: 9909
Score: 80.00 Matches: 35
Percent Similarity: 40.37% Conservative: 30
Best Local Similarity: 21.74% Mismatches: 72
Query Match: 7.43% Indels: 24
DB: 4 Gaps: 7

US-10-029-020-14_COPY_2400_2600 (1-201) x US-08-961-527-12 (1-9909)

QY 59 IleSerAsn---SerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeu--- 76
DB 3313 GTTTCCAATGTTAGCATGAGTTACGGACTCTCTGCTAGCTAGCTAAATCCTATCTTGA 3254
QY 77 ---LeuThrPheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAspMet 95
DB 3253 GCCTGGATGAGGGCGCTTTGTGTGAAACTGTAGCACCAGACTTATCAAGGTTTCTCTT 3194
QY 96 AspAlaMetGluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAsp 115
DB 3193 GATGAGACCAACCGTATGATGCGCATGGTGGAGGATCTCTCCATCTTTCAGTATTGAT 3134
QY 116 AsnSerLysSerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheVal 135
DB 3133 AATGCTACAGTCACCTAGATGTGGAACGTG-----ATTAACTTCACGTCTTAT 3083
QY 136 Thr-----LeuGluArgPheAspGlnLeuTyrGlySer----- 146

```
Db 3082 ACCTTTATCTCAATCTGTTTGCACAGATCAAGACGACGAAAGAGAGAAAAATATGAG 3023
Qy 147 -----ThrIleThrSerCysGlnGlnAlaProLysThrLysLysPheAla 161
Db 3022 TTGGTGAGAGATTATCCCAATCTTATCTGGATGGAAATTTGATACAGATAAGATGACG 2963
Qy 162 Ser----SerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 2962 CAGGTTCTGCACATATTTAAATTAATGCTATTAAATGCTATTCGCACAGATGGGGGTAAATC 2903
Qy 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
Db 2902 ACT-----GTCAGATGAGACAACTGAACACAGATGATTTTATCCATTTCTGAC 2852
Qy 201 His 201
Db 2851 CAC 2849

RESULT 9
US-09-620-312D-86
; Sequence 86, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 86
; LENGTH: 3215
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (264)..(3128)
US-09-620-312D-86

Alignment Scores:
Pred. No.: 7.99 Length: 3215
Score: 78.50 Matches: 48
Percent Similarity: 35.53% Conservative: 33
Best Local Similarity: 21.05% Mismatches: 74
Query Match: 7.29% Indels: 73
DB: 4 Gaps: 10

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-620-312D-86 (1-3215)

Qy 1 IlesGlyThrHisGlyGlyLeu-----TyrAsp 9
Db 1017 ATTGGGTACAATGGGAACCTCAITTCAGTGTGTTAAGACCCCTGCCCAAGAAATTATGAT 1076
```



```
RESULT 14
US-09-717-926-1
; Sequence 1, Application US/0917926
; Patent No. 6569657
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Cook, William J.
; TITLE OF INVENTION: 32140, A No. 6569657el Human Aldehyde
; FILE OF INVENTION: Dehydrogenase
; FILE REFERENCE: 35800/205243
; CURRENT APPLICATION NUMBER: US/09/717,926
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 7220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (129)...(2900)
US-09-717-926-1
Alignment Scores:
Pred. No.: 52.2 Length: 7220
Score: 76.50 Matches: 31
Percent Similarity: 38.8% Conservative: 18
Best Local Similarity: 24.6% Mismatches: 48
Query Match: 7.1% Indels: 29
DB: 4 Gaps: 6
US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-717-926-1 (1-7220)
QY 88 ProGlyTyrProLysProAspMetAspAlaMetGluProSerTyrGluLeuIleHisThr 107
Db 759 CPTCGTATATACCCAGCAGAGGAGGCA-----ACATATGAAGGTATCCAGAA 809
QY 108 GlnMetLysThrGln---GluTrpAspAsnSerLysSerIleLeu----- 121
Db 810 AAGAAATGCTCGAGATTCTTGGGACCACTGCGCGAGTTTACATACTGGATTGCA 869
QY 122 -----GlyValGlnCysGluValGlnLysGlnLysAlaPheVal 135
Db 870 GGTCAATGATAAAGTCCCTGGAGCTTGCACAGATAAATGACAGATGTCATTTC--- 926
QY 136 ThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaPro 155
Db 927 -----TATGGCTCGACATTACTGAATAGCTCTGTCGCTCCT 962
QY 156 -----LysThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyVal 171
Db 963 GGAGAACCACTGGAATTAAGTGCCAGAGCCCTGCTCGTTACCAAAATGGACTT 1022
QY 172 LysPheAlaLeuLysAspGlyArgValThrThrAspIleIleSerValAlaAsnGluAsp 191
Db 1023 GTTCTTTTGGTAACGATGGAAGAGCACTACG---GTGAGAAATCTGAGTTTGAAGAT 1079
QY 192 GlyArgArgValAlaAla 197
Db 1080 GGAAAAATGATCCCTGCC 1097
RESULT 15
US-09-817-514A-5
; Sequence 5, Application US/09817514A
; Patent No. 6639129
; GENERAL INFORMATION:
; APPLICANT: ffrench-Constant, Richard
; APPLICANT: Bowen, David
; APPLICANT: Rocheleau, Thomas
; APPLICANT: Waterfield, Nicholas
; TITLE OF INVENTION: DNA SEQUENCES FROM PHOTORHABDUS LUMINESCENS
```

```
; FILE REFERENCE: 61645
; CURRENT APPLICATION NUMBER: US/09/817,514A
; CURRENT FILING DATE: 2000-03-26
; PRIOR APPLICATION NUMBER: US 60/191806
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 2745
; TYPE: DNA
; ORGANISM: Photorhabdus luminescens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2745)
US-09-817-514A-5
Alignment Scores:
Pred. No.: 13.4 Length: 2745
Score: 76.00 Matches: 22
Percent Similarity: 43.33% Conservative: 4
Best Local Similarity: 36.67% Mismatches: 26
Query Match: 7.06% Indels: 8
DB: 4 Gaps: 2
US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-817-514A-5 (1-2745)
QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 1816 ATTGCTATTCCGCGCAAGAACAGAT---GCCACCGGGTTGTATTATTACGGTTATCGT 1872
QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 1873 TATTACCACACCGTGGCGGCAGATGTTAAGCGCGAC----- 1911
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
Db 1912 CCGCAGGAACCAATTGATGGGCTGAATCTATACCGAATGTTAAGAAATAATCTGTGAGT 1971
Search completed: August 14, 2004, 21:40:27
Job time : 121.466 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 2357.29 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_2400_2600
Perfect score: 1077
Sequence: 1 IGYHGLYDLTKLVHMR.....TDIISVANEDGRRVAALNH 201

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/cpn2.1/USPTO.spool/US10029020/runat_06082004.112216.29287/app_query.fasta_1.3519
-DB=EST -QFMR=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITG=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CEN 1.1 13135 @runat_06082004.112216.29287 -NCPV=6 -ICPV=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vrt: *
21: em_gss_fun: *
22: em_gss_mam: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_rod: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gss1: *

29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Length | DB | ID | Description |
|------------|-------|---------|--------|----|----------|---------------------|
| 1 | 1077 | 100.0 | 6246 | 29 | AY413475 | AY413475 Homo sapi |
| 2 | 1071 | 99.4 | 5970 | 29 | AY413476 | AY413476 Pan trogl |
| 3 | 1030 | 95.6 | 668 | 13 | BQ563239 | BQ563239 GI01b09.Y |
| 4 | 951 | 88.3 | 734 | 13 | BU233488 | BU233488 603408464 |
| 5 | 896 | 83.2 | 637 | 14 | CF171639 | CF171639 B0845C10- |
| 6 | 836 | 77.6 | 667 | 10 | AW767437 | AW767437 GA69B01.Y |
| 7 | 816 | 75.8 | 850 | 13 | BU232300 | BU232300 603341548 |
| 8 | 700 | 65.0 | 718 | 9 | AL045768 | AL045768 DKEZP434F |
| 9 | 691 | 64.2 | 525 | 9 | AL046228 | AL046228 DKEZP434E |
| 10 | 642.5 | 59.7 | 847 | 13 | BU234988 | BU234988 603408474 |
| 11 | 636.5 | 59.1 | 3038 | 11 | AK037897 | AK037897 Mus muscu |
| 12 | 631.5 | 58.6 | 929 | 13 | BU371366 | BU371366 603598013 |
| 13 | 629 | 58.4 | 506 | 10 | BE819779 | BE819779 MK3-BN036 |
| 14 | 618.5 | 57.4 | 621 | 14 | CB578819 | CB578819 AMGNNUC:N |
| 15 | 566 | 52.6 | 776 | 14 | CA777388 | CA777388 ip19c06.x |
| 16 | 566 | 52.6 | 5094 | 29 | AY405420 | AY405420 Homo sapi |
| 17 | 552.5 | 51.3 | 616 | 13 | BX506934 | BX506934 DKEZP7791 |
| 18 | 527.5 | 49.0 | 521 | 14 | CB720876 | CB720876 AMGNNUC:N |
| 19 | 511.5 | 47.5 | 885 | 13 | BU120313 | BU120313 603142826 |
| 20 | 502 | 46.6 | 555 | 12 | BG732516 | BG732516 333252 MA |
| 21 | 498.5 | 46.3 | 728 | 14 | CA344273 | CA344273 674645 NC |
| 22 | 498 | 46.2 | 720 | 13 | BU346142 | BU346142 603525314 |
| 23 | 484 | 44.9 | 428 | 14 | CB794379 | CB794379 AMGNNUC:N |
| 24 | 470 | 43.6 | 5069 | 29 | AY405421 | AY405421 Pan trogl |
| 25 | 466 | 43.3 | 625 | 10 | BE865229 | BE865229 UI-M-BH2. |
| 26 | 456 | 42.3 | 748 | 9 | AUI69909 | AUI69909 AUI69909 |
| 27 | 443.5 | 41.2 | 678 | 13 | BQ572906 | BQ572906 UI-M-FD0- |
| 28 | 438.5 | 40.7 | 625 | 13 | BX308610 | BX308610 BX308610 |
| 29 | 438.5 | 40.7 | 732 | 14 | CB519165 | CB519165 UI-M-GH0- |
| 30 | 436.5 | 40.5 | 624 | 13 | BX308609 | BX308609 BX308609 |
| 31 | 429.5 | 39.9 | 5087 | 29 | AY405422 | AY405422 Mus muscu |
| 32 | 424 | 39.4 | 621 | 12 | BI400254 | BI400254 MI-P-A11- |
| 33 | 421.5 | 39.1 | 731 | 14 | CF539420 | CF539420 UI-M-GH0- |
| 34 | 418.5 | 38.9 | 654 | 12 | BM490220 | BM490220 B9P2n.pK0 |
| 35 | 414 | 38.4 | 818 | 14 | CD807778 | CD807778 UI-M-GW0- |
| 36 | 413.5 | 38.4 | 446 | 14 | CB545337 | CB545337 AMGNNUC:N |
| 37 | 405.5 | 37.7 | 603 | 13 | BX300081 | BX300081 BX300081 |
| 38 | 397 | 36.9 | 473 | 14 | CD736507 | CD736507 4019565.1 |
| 39 | 388.5 | 36.1 | 564 | 9 | AI820014 | AI820014 wj60407.x |
| 40 | 386.5 | 35.9 | 650 | 10 | BB576857 | BB576857 BB576857 |
| 41 | 386.5 | 35.9 | 3760 | 11 | AK031268 | AK031268 Mus muscu |
| 42 | 381.5 | 35.4 | 705 | 9 | AA142988 | AA142988 xl48c01.r |
| 43 | 380 | 35.3 | 232 | 14 | T05481 | T05481 EST033370 Fe |
| 44 | 378.5 | 35.1 | 610 | 14 | CB580731 | CB580731 AMGNNUC:N |
| 45 | 375.5 | 34.9 | 851 | 14 | CB179955 | CB179955 AGENCOURT |

ALIGNMENTS

RESULT 1
AY413475
LOCUS
DEFINITION Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY413475.1 GI:39769437
VERSION AY413475.1
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 6246)

181 ThrThrAspIleIleSerValAlaAenGluAspGlyArgValAlaAlaIleLeuAsn 200
 5674 ACCACGACATCATCATGCTGGCCATGAGATGGCGAAGGTTGCTGCTCATCTTGAAC 5733

201 His 201
 5734 CAT 5736

RESULT 2
 AY413476 5970 bp DNA linear GSS 12-DEC-2003
 LOCUS Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY413476
 VERSION AY413476.1 GI:39769438
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 REFERENCE 1 (bases 1 to 5970)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 5970)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES source
 1..5970
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>5970
 /locus_tag="HCM4903"

gene

ORIGIN

Alignment Scores:
 Pred. No.: 6,198-126 Length: 6246
 Score: 1077.00 Matches: 201
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 29 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AY413475 (1-6246)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
 DB 5134 ATAGGCTACCATGCTGGCTCTATGATCCACCAAGCTTGTCCACATGGCGCGCA 5193

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
 DB 5194 GATTATGATGCTGGCGGAGCTGGCTAGCTAGCCAGACACGAGCTGTGGAAGCACCTT 5253

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
 DB 5254 AGTAGCAGCAACGTCATGCTTTTAACTCTATATCTTCAAAAACAAACCCCATCAGC 5313

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
 DB 5314 AACTCCAGGACATCAAGTGTTCATGACAGATGTTAAACAGCTGGCTGTCTACCTTTGA 5373

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 DB 5374 TTCCAGCTACACACGTCATGCTTATCCCAACACAGACATGATGCCATGGAACCC 5433

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
 DB 5434 TCCTACGAGCTCATCCACACACAGATGAAACAGCAGGAGTGGACACAGCAAGTCTATC 5493

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 DB 5494 CTGGGGTACAGTGTGAAGTACAGACAGCTCAAGCCCTTGTCTACCTTAGAACGTTT 5553

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
 DB 5554 GACAGCTCATGGCTCCACATCACAGCTGCAGCAGGCTCCAAAGACCAAGAGTTT 5613

QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
 DB 5614 GCATCCAGCGCTCAGTCTTTGGCAAGGGGTCAAGTTTCCCTTGAAGGATGGCCGAGTG 5673

181 ThrThrAspIleIleSerValAlaAenGluAspGlyArgValAlaAlaIleLeuAsn 200
 5674 ACCACGACATCATCATGCTGGCCATGAGATGGCGAAGGTTGCTGCTCATCTTGAAC 5733

201 His 201
 5734 CAT 5736

RESULT 2
 AY413476 5970 bp DNA linear GSS 12-DEC-2003
 LOCUS Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY413476
 VERSION AY413476.1 GI:39769438
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 REFERENCE 1 (bases 1 to 5970)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 5970)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES source
 1..5970
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>5970
 /locus_tag="HCM4903"

gene

ORIGIN

Alignment Scores:
 Pred. No.: 6,198-126 Length: 6246
 Score: 1077.00 Matches: 201
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 29 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AY413475 (1-6246)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
 DB 5134 ATAGGCTACCATGCTGGCTCTATGATCCACCAAGCTTGTCCACATGGCGCGCA 5193

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
 DB 5194 GATTATGATGCTGGCGGAGCTGGCTAGCTAGCCAGACACGAGCTGTGGAAGCACCTT 5253

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
 DB 5254 AGTAGCAGCAACGTCATGCTTTTAACTCTATATCTTCAAAAACAAACCCCATCAGC 5313

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
 DB 5314 AACTCCAGGACATCAAGTGTTCATGACAGATGTTAAACAGCTGGCTGTCTACCTTTGA 5373

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 DB 5374 TTCCAGCTACACACGTCATGCTTATCCCAACACAGACATGATGCCATGGAACCC 5433

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
 DB 5434 TCCTACGAGCTCATCCACACACAGATGAAACAGCAGGAGTGGACACAGCAAGTCTATC 5493

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 DB 5494 CTGGGGTACAGTGTGAAGTACAGACAGCTCAAGCCCTTGTCTACCTTAGAACGTTT 5553

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
 DB 5554 GACAGCTCATGGCTCCACATCACAGCTGCAGCAGGCTCCAAAGACCAAGAGTTT 5613

QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
 DB 5614 GCATCCAGCGCTCAGTCTTTGGCAAGGGGTCAAGTTTCCCTTGAAGGATGGCCGAGTG 5673

```

QY      81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db      5098 TTCCAGCTACACACGTGATCCCTGTTATCCCAACACGACATGATGCCATGGAAACC 5157

QY      101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db      5158 TCTACGAGCTCATCCACACACAGATGAAACGACGAGTGGGACACACGCAAGTCTATC 5217

QY      121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db      5218 CTGGGTTACAGTGTGAAGTACAGACGAGCTCAAGGCCCTTTGTACCTTAGAACGGTTT 5277

QY      141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db      5278 GACCAGCTCTATGGTCCACATCATCACCTGCCAGCGGCTCCAAAGACCAAGAGTTT 5337

QY      161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db      5338 GCATCCAGCAGCTCAGTCTTTGGCAAGGGGGTCAAGTTTGCCTTGAAGGATGCCGAGTG 5397

QY      181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn 200
Db      5398 ACCACAGACATCATGATGTGCCAATGAGGATGGCGAAGGGTTGCTGCCATCTTGAAC 5457

QY      201 His 201
Db      5458 CAT 5460

RESULT 3
B0563239
LOCUS      B0563239          668 bp      mRNA      linear      EST 19-JUN-2002
DEFINITION clone gi01b09 5', mRNA sequence.
ACCESSION B0563239
VERSION   B0563239.1  GI:21466220
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
TITLE     1 (bases 1 to 668)
KACHAR, B.
EST analysis of gene expression in the mouse Organ of Corti at the
onset of hearing
Unpublished (2002)
Contact: Kachar, B.
Structural Cell Biology
National Institute of Deafness and other Communication Disorders
50/4249 South Drive, NIH, Bethesda, MD 20892-8027, USA
Tel: 301-402-1599
Fax: 301-402-1765
Email: kachar@nidcd.nih.gov
Plate: 01 row: b column: 09
Seq primer: M13Rpl reverse primer (ABI).
Location/Qualifiers
    1..668
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="BALE/c"
    /db_xref="taxon:10090"
    /clone="gi01b09"
    /sex="male and female"
    /dev_stage="Post natal day 5 to 13"
    /note="Organ: Organ of Corti; Vector: pBluescript; The
organ of Corti (OC) was fine dissected from a total of 386
OC as follows: 102 samples from post-natal (P) day 5; 72
from P6; 60 from P7; 46 from P8; 18 from P9; 20 from P10;
14 from P12 and 24 from P13. After killing animals by
cervical dislocation followed by decapitation, the bulla
was removed and opened in Leibowitz medium. The bony
capsule of the cochlea was chipped away, stria vascularis

```

and spiral ligament were removed and the sensory epithelium was carefully dissected out of the modiolus. Total RNA was extracted using the micro Fasttrack kit (catalog # K1593-02; Invitrogen, Carlsbad, CA), according to manufacturer's instructions. Reverse transcription and library construction were carried out with the Uni-Zap XR vector kit (catalog # 237211, Stratagene) and Uni-Zap XR Gigapack III Gold Cloning kit (catalog # 237612), both from Stratagene (La Jolla, CA, USA), according to manufacturer's instructions. Briefly: 1.5 ug mRNA was reverse transcribed using a hybrid oligo(dT) linker-primer that contains an Xho I site. First strand synthesis was primed with the linker-primer and transcribed using Moloney murine leukemia virus reverse transcriptase (MMLV-RT) and 5-methyl dCTP. The second strand was synthesized with DNA polymerase and RNase H. Complementary DNA was blunt ended with Pfu DNA polymerase, ligated with EcoR I adapters in the presence of ligase and digested with Xho I. The cDNA was sequentially size fractionated over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden) and Clontech Chroma Spin-1000 (Clontech, Palo Alto, CA) columns to enrich for cDNAs greater than 400bp and 1000 bp, respectively. The cDNA was then directionally ligated to the Uni-ZAP XR vector, which had been predigested with EcoR I and Xho I. The phagemid was packaged with Gigapack III Gold and, upon titration on XL1 Blue MRF' cells, the yield of the phage library was estimated to be 11,100,000 recombinants. Stratagene's ExAssist interference resistance helper phage (catalogue # 211203) was adopted to rescue plasmid DNA from the phages. Upon plating of the rescued library, individual cDNA clones were selected and grown in 96-well, 2 ml growth plate. Plasmid DNA was purified from 200 ul of saturated culture with the Concert96(TM) plasmid purification kit (Invitrogen, Carlsbad, CA) as instructed by the manufacturer. ESTs from the 5' end of the cDNA clones were generated with the universal Big3 reverse primer (CAGGAACAGCTAGTACC) and 25x strength Big3 terminator sequencing chemistry (Applied Biosystems, Foster City, CA). Sequencing reactions were performed on MJ Tetrad thermal cyclers (MJ Research, Waltham, MA), and analyzed on 3700 automated capillary sequencers using POP5 polymer (Applied Biosystems, Foster City, CA). The frequency distribution of the library is as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10; 1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of genes are present in GenBank and have known function; 23% have hits in GenBank, but do not have assigned function; 12% are uncharacterized ESTs and 20% are unidentified."

ORIGIN

```

Alignment Scores:
Pred. No.:      2e-121      Length:      668
Score:          1030.00     Matches:    188
Percent Similarity: 99.00%   Conservative: 11
Best Local Similarity: 93.53% Mismatches:    2
Query Match:      95.64%    Indels:      0
DB:               13       Gaps:        0

```

US-10-029-020-14_COPY_2400_2600 (1-201) x BQ563239 (1-668)

```

QY      1 IleGlyTyrHisGlyCysLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db      1 GTCCGGTACACGGGGCGCTCTATGATCCACTCACCAGCTTGTCCACATGGCCGACCG 60

QY      21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTyrPheHisLeu 40
Db      61 GATTATGATGTCTGGCTGGACGCTGGCAACCCAGACCATGAACTCTGGAACGCGCTG 120

QY      41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db      121 AGTACCAACAGCATCGTCCCTTTTCATCTCTACATGTTTAAAGAACCAACCCCATCAGC 180

QY      61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80

```

methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores: 3.31e-111 Length: 734
Pred. No.: 951.00 Matches: 174
Score: 951.00 Conserv: 13
Percent Similarity: 94.92% Mismatches: 10
Best Local Similarity: 88.32% Indels: 1
Query Match: 88.30% Gaps: 0
DB: 13

US-10-029-020-14_COPY_2400_2600 (1-201) x BU233488 (1-734)

QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 92 ATCGGCTACCATGGAGGACTGATGATCCCTCACAAGCTCATCCACATGGACGGAGA 151
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuThrLysHisLeu 40
Db 152 GACTATGATGCTGGCAGTGGTGACAGCTCCAGACCATGATATGGAGACACCTG 211
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProLysSer 60
Db 212 AGTAGCATAACATCATGCTTTCACCTGATATGTTTCAAAACAATATCCATCAGC 271
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 272 AACTCTCAGATATCAATGCTACATGACAGATGTCAACAGTTGGTACTCAC-TTTGGG 330
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 331 TTCCAGCTACACAATGCTATCCCTGGATACCCCAAGCCAGACCTGGATGCCATGGAGCCA 390
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnLutrpAspAsnSerLysSerIle 120
Db 391 TCATATGAGCTTATCACACGACGAGATGAAACCCAGATGGACACACCAAGTCAATT 450
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 451 TTAGGGGTCCCAATGTGAAGTGCAGAGCAGCTGAAGGCCCTTTGTCACTCTTGAGCGCTTT 510
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 511 GAACAGATCTACAGCTCCAGATCCCGGGTGCACACAGGTCACAGAGAACCAAGATTC 570
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 571 GCCTCCGGAGGGTCTATCTTCGCAAGGTGTCAAGTTTCCATGAGAGGAGCGCGTGT 630
QY 181 ThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAla 197
Db 631 GCCACGACATCATCAGTGTGGCCATGAGGACGGCGGAGGATCGCAGCA 681

RESULT 5

CF171639 637 bp mRNA linear EST 25-JUL-2003
LOCUS B0845C10-5 NIA Mouse Newborn Kidney cDNA Library (Long 1) Mus
DEFINITION musculus cDNA clone NIA:B0845C10 IMAGE:30471777 5', mRNA sequence.
ACCESSION CF171639 GI:33281188
VERSION CF171639.1
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Db 181 AACTCTCAGGACATCAAGTGTCTTCATGACAGATGTCAACAGCTGGCTCCTCACCTTTGGA 240
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 241 TTCCAGCTGGCAACAGTGTATACCTGGCTATCCCAAGCCAGACACAGATGCGCATGCCACCC 300
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
Db 301 TCCTACGAGCTCGTACACACAGATGAAACTCAGGAATGGCAACAGCAAGTCTATC 360
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 361 CTCGGGGTACAGTGTGAAGTTTCAAGAACAACTCAAGGCTTTGCTTACCTTAGAAGCTTT 420
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 421 GACCAAGCTCTACGGCTCGACCATCACCAGCTGCCAACAGAGGCCCTTGAGACAAAGAAGTTT 480
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 481 GCCTCCAGTGTTCATCTTTGGCAAGGGGCTCAAGTTTGCCTTGAAGATGGTGGAGTG 540
QY 181 ThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
Db 541 ACCACTGACATCATCATGCTGCGCAATGAGGATGGCGGAGGATCGCAGCCATCTTGAAC 600
QY 201 His 201
Db 601 AAT 603

BU233488 734 bp mRNA linear EST 26-NOV-2002
LOCUS 603408464F1 CSFQCHN24 Gallus gallus cDNA clone CHEST321p15 5', mRNA
DEFINITION sequence.
ACCESSION BU233488
VERSION BU233488.1 GI:25477142
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 734)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Rong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..734
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, HiseX"
/db_xref="taxon:9031"
/clone="CHEST321p15"
/dev_stage="22"
/lab_host="DH10B"
/clone_lib="CSFQCHN24"
/note="Organ: heads; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using

REFERENCE
 1. (bases 1 to 637)
 AUTHORS
 TITLE
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method

JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 Genome Res. 11 (9), 1553-1558 (2001)
 21429098
 11544199
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cda@lgsun.grc.nia.nih.gov
 Plate: B0845 row: C column: 10
 Seq primer: M13 Reverse
 High quality sequence stop: 637
 POLYA=No.

FEATURES
 source
 Location/Qualifiers
 1..637
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="niaEST:B0845C10-5"
 /db_xref="taxon:10090"
 /clone="NIA:B0845C10 IMAGE:30471777"
 /dev_stage="Newborn Kidney"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long
 1)"

/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
 Site 2: NotI; Mouse cDNA project by the Laboratory of
 Genetics, National Institute on Aging (NIA), Intramural
 Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
 In brief, double-stranded cDNAs were synthesized with an
 oligo(dT) primer [Invitrogen:
 5'-PGACTAGTTCAGATCGGCGCCCTTTT-3'] from
 26 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to lone-linker LI-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
 vector. The DH10B E. coli host was transformed with the
 ligation mixture by the standard chemical method. The
 average insert size is about 3.0 kb. The library was
 constructed by Yulan Piao."

ORIGIN
 Alignment Scores:
 Pred. No.: 3,11e-104 Length: 637
 Score: 896.00 Matches: 162
 Percent Similarity: 98.84% Conservative: 8
 Best Local Similarity: 84.19% Mismatches: 2
 Query Match: 93.19% Indels: 0
 DB: 14 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x CF171639 (1-637)

QY 1 IleGlyTyHisGlyGlyLeuTyAspProLeuThrLysLeuValHisMetGlyArgG 20
 Db 120 ATCGGTACACGGCGGCTCTATGATCCACTCACCAAGCTGTCCAGTGGCCGACGG 179

QY 21 AspTyAspValLeuAlaGlyArgTrpThrSerProAspHisGlyLeuTrpLysHisLeu 40
 Db 180 GATTATGATGCTGCTGGCGCTGGCAAGCCAGACCATGAACTCTGGAAACGCCTG 239

QY 41 SerSerSerAsnValMetProPheAsnLeuTyMetPheLysAsnAsnProIleSer 60
 Db 240 AGTACACACAGCATCGTGCCTTTTCATCTCTCATGTTTAAAGACAAACCAACCATCAGC 299

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
 Db 300 AACTCTCAGGACATCAAGTGCTTCATGACAGATGTCACAGCTGGCTCTCCTTGTGA 359

QY 81 PheGlnLeuHisAsnValIleProGlyTyProLysProAspMetAspAlaMetGluPro 100
 Db 360 TTCCAGCTGCACACAGCTGATACCTGGCTATCCCAAGCCACAGACATGCCATGGAACCC 419

QY 101 SerTyGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
 Db 420 TCTTACGAGCTCGTACACACAGATGAATACTCAGGAATGGGACACACAGCAAGTCTATC 479

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 Db 480 CTCGGGTACAGTGTGAAGTTTCAGAACTCAAGGCTTCGTTACCTTGAACGCTTT 539

QY 141 AspGlnLeuTyGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
 Db 540 GACAGCTCTACGGCTCGACCATCACAGCTGCCAACAGGCGCTGAGACAAAGAGTTT 599

QY 161 AlaSerSerGlySerValPheGlyLysGlyValLys 172
 Db 600 GCCTCCAGTGGTTCATCTTTGGCAAGGGGTCAAG 635

RESULT 6
 AW767437
 LOCUS
 DEFINITION
 AW767437 667 bp mRNA linear EST 16-FEB-2001
 da69b01.yl Harland stage 19-23 Xenopus laevis cDNA clone
 IMAGE:3200137 5' similar to TR:070465 070465 DOC4. ; mRNA
 sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 AW767437.1 GI:7699442
 EST
 Xenopus laevis (African clawed frog)

REFERENCE
 AUTHORS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 667)

TITLE
 JOURNAL
 COMMENT
 Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L.,
 Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y.,
 Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,
 Waterston, R. and Willson, R.
 WashU Xenopus EST project, 1999
 Unpublished (1999)
 Contact: Sandy Clifton, Ph.D.
 WashU Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed by R. Harland, PhD. (University of California,
 Berkeley)

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at:
 image.llnl.gov/image/html/resources.shtml
 Seq primer: -40RP from Gibco

High quality sequence stop: 494.
 Location/Qualifiers
 1..667
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:3200137"
 /tissue_type="neurula"
 /dev_stage="stage 19-23"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Harland stage 19-23"
 /note="Vector: PCS107 (custom); Site 1: NotI; Site 2:
 SalI; cDNA made by oligo-dT priming. Library constructed


```

QY 131 LeuLysAlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSer 150
Db 307 CTGAAGCGCTTTGTCATCTTGAGCGCTTTGAACAGATCTACAGCTCCAGCATCGCGGG 366
QY 151 CysGlnGlnAlaProLysThrLysLysPheAlaSerSerGlySerValPheGlyLysGly 170
Db 367 TGCCACACAGTCAAGAAGAACAGAACTTTGGCCCTCCGAGGGTCTATCTTCGGCAAGGT 426
QY 171 ValLysPheAlaLeuLysAspGlyArgValThrThrAspIleIleSerValAlaAsnGlu 190
Db 427 GTCAAGTTTGCCATGAAGCAGCGGCGTGTGCCACCCACATCATCATGTGTGGCCAATGAG 486
QY 191 AspGlyArgValAlaAlaIleLeuAsnHis 201
Db 487 GACGGGGAGGATCGCAGCCATCTTGAACAAT 519

RESULT 8
LOCUS AL045768
DEFINITION DKFZp343F206_r1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045768
VERSION AL045768.1 GI:5433880
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 718)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Koehrer, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
sl sequence also available.
This clone (DKFZp343F206) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..718
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp343F206"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Alignment Scores:
Pred. No.: 5,89e-79 Length: 718
Score: 700.00 Matches: 134
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.00% Indels: 0
Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AL045768 (1-718)

QY 68 PheMetThrAspValAsnSerTrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIle 87
Db 3 TTCATGACAGATGTTAAACAGCTGGTGTCTACCTTTGGATTCCAGCTACAAACGTGATC 62
QY 88 ProGlyTyrProLysProAspMetAspAlaMetGluProSerTyrGluLeuIleHisThr 107

```

```

Db 63 CTGTGTTATCCCAAAACAGACATGGATGCCATGGAACCTCTCCAGAGCTTATCCACA 122
QY 108 GlnMetLysThrGlnGluTrpAspAsnSerLysSerIleLeuGlyValGlnCysGluVal 127
Db 123 CAGATGAACACGACAGAGTGGGACAAACAGCAAGTCTATCTCTGGGTACAGTGTGAAGTA 182
QY 128 GlnLysGlnLeuLysAlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThr 147
Db 183 CAGAAGCAGCTCAAGGCCCTTTGTACCTTAGAACGGTTTGACCAGCTCTATGGCTCCACA 242
QY 148 IleThrSerCysGlnGlnAlaProLysThrLysLysPheAlaSerSerGlySerValPhe 167
Db 243 ATCACCAGCTGCCAGCAGGCTCCAAAGACCAAGAGTTTGCAATCCAGCGCTCAGTCTTT 302
QY 168 GlyLysGlyValLysPheAlaLeuLysAspGlyArgValThrThrAspIleIleSerVal 187
Db 303 GGCAAGGGGGTCAAGTTTGCTTGAAGGATGCCGAGTGACCACAGACATCATCATGTGTG 362
QY 188 AlaAsnGluAspGlyArgValArgValAlaAlaIleLeuAsnHis 201
Db 363 GCCAATGAGGATGGCGGAGGGTGTCTGCCATCTTTGAACCAT 404

RESULT 9
LOCUS AL046228
DEFINITION DKFZp343E177_r1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL046228
VERSION AL046228.1 GI:5434312
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 525)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Koehrer, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp343E177) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..525
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp343E177"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Alignment Scores:
Pred. No.: 5,15e-78 Length: 525
Score: 691.00 Matches: 133
Percent Similarity: 99.25% Conservative: 0
Best Local Similarity: 99.25% Mismatches: 1
Query Match: 64.16% Indels: 0
Gaps: 0

```

Mon Aug 16 09:01:13 2004

US-10-029-020-14_COPY_2400_2600 (1-201) x AL046228 (1-525)

QY 68 PheMetThrAspValAsnSerTrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIle 87
 Db 3 TTCATGACAGATGTTAAACAGCTGGCTGCTACCGGGGATTCACAGCTACACAGCTGATC 62
 QY 88 ProGlyTyrProLysProAspMetAspAlaMetGluProSerTyrGluLeuIleHisThr 107
 Db 63 CTTGTTATCCAAACACAGACATGATGCCATGAAACCTCTACGAGCTTATCCACACA 122
 QY 108 GlnMetLysThrGlnGluTrpAspAsnSerLysSerIleLeuValGlnCysGluVal 127
 Db 123 CAGATGAAACGACAGGAGTGGACACACAGAGCTTATCTTCGGGGTACAGTGTGAAGTA 182
 QY 128 GlnLysGlnLeuLysAlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThr 147
 Db 183 CAGAACGACGCTCAGGCTTTGTACCTTGAAGGATGCGGAGTGCACACACATCATCATG 242
 QY 148 IleThrSerCysGlnGlnAlaProLysThrLysLysPheAlaSerSerGlySerValPhe 167
 Db 243 ATCACCAGCTGCAGCAGGCTCCAAAGACCAAGAGTTTGCATCCAGCGGCTCAGTCTTT 302
 QY 168 GlyLysGlyValLysPheAlaLeuLysAspGlyArgValThrThrAspIleIleSerVal 187
 Db 303 GGCAGAGGGGTCACAGTTTCCCTTGAAGGATGCGGAGTGCACACACATCATCATG 362
 QY 188 AlaAsnGluAspGlyArgValAlaAlaIleLeuAsnHis 201
 Db 363 GCAATGAGGATGGCGCAAGGTTGTGTCACCTTGAACCAT 404

RESULT 10
 BU234988 847 bp mRNA linear EST 26-NOV-2002
 LOCUS 603408474F1 CSEQCHN24 Gallus gallus cdna clone CHEST321d19 5', mRNA
 DEFINITION sequence.

ACCESSION BU234988
 VERSION BU234988.1 GI:25479972
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 847)
 AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
 TITLE A Comprehensive Collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 22335534
 PUBMED 12445392

COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 0161208930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers

FEATURES
 source 1..847
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, HiseX"
 /db_xref="taxon:9031"
 /clone="CHEST321d19"
 /dev_stage="22"
 /lab_host="DH108"
 /clone_lib="CSEQCHN24"
 /note="Organ: heads; Vector: pBluescript II KS(+); Site_1:
 EcoRI; Site_2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores: 1.88e-71 Length: 847
 Pred. No.: 642.50 Matches: 118
 Score: 77.11% Conservative: 37
 Percent Similarity: 77.11% Mismatches: 45
 Best Local Similarity: 58.71% Indels: 2
 Query Match: 59.66% Gaps: 1
 DB: 13

US-10-029-020-14_COPY_2400_2600 (1-201) x BU234988 (1-847)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
 Db 141 ATTGTTTTCATGGAGGCTCTATGATCTCTCACAAATTTGTTTCATCTGGGTGACGG 200
 QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
 Db 201 GATTATGATGTTATGTCGTGGTGGACACCAACCATCATCATATGGAACACCTG 260
 QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
 Db 261 AATGCTGTC---CCCAACCATTCATCTCTACTCATTTGAAATAACTACCCAGTTGGC 317
 QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
 Db 318 AGGATCCAAGATGTTGCTTAAGTATACACAGACATTTGGAAGTTGGGTAGAGCTATT 377
 QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 Db 378 TTCAGTTGCACAAATGACTACTCCTGGTTCCTCAAAACCCAGAAATCGAAGCTTG-GAGACA 436
 QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
 Db 437 ACATATGAATCTTCTACAGCTTCAACAAACCCAGAGTGGGATCCTCGAAAGACTATC 496
 QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 Db 497 CTGTGTTATTCAGTGTGAGCTACAGAGCAACTCCGAAACTTTATCTCTGGATCAACTT 556
 QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
 Db 557 CCAATGACCCCGAGGTATAGTATGGCAAGTGTATGAGGAGTGAGGAGCAACCCAGGTTT 616
 QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
 Db 617 GCAGCTATTCTCTCAGTATTGGAAGGATCAAAATTTGCTATCAAGGATGGCATCGTC 676
 QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
 Db 677 ACAGCGGACATTTATTGTTGGTGTGCTAATGAGGACACCGCGGCATCGCTGCTACTAC 736
 QY 201 His 201
 Db 737 AAT 739

RESULT 11
 AK037897
 LOCUS
 DEFINITION Mus musculus 16 days neonate thymus cDNA, RIKEN full-length
 enriched library, clone:AI30060L24 product:odd Oz/ten-m homolog 1
 (Drosophila), full insert sequence.
 ACCESSION AK037897
 VERSION AK037897.1 GI:26332325

AK037897 3038 bp mRNA linear HTC 19-SEP-2003
 Mus musculus 16 days neonate thymus cDNA, RIKEN full-length
 enriched library, clone:AI30060L24 product:odd Oz/ten-m homolog 1
 (Drosophila), full insert sequence.

| | | | |
|-----------|---|--------|---------|
| KEYWORDS | HTC; CAP trapper. | source | 1. 3038 |
| SOURCE | Mus musculus (house mouse) | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| REFERENCE | Carninci, P. and Hayashizaki, Y. | | |
| AUTHORS | | | |
| TITLE | High-efficiency full-length cDNA cloning | | |
| JOURNAL | Meth. Enzymol. 303, 19-44 (1999) | | |
| MEDLINE | 99279253 | | |
| PUBMED | 10349636 | | |
| REFERENCE | | | |
| AUTHORS | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | | |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | | |
| JOURNAL | Genome Res. 10 (10), 1617-1630 (2000) | | |
| MEDLINE | 20499374 | | |
| PUBMED | 11042159 | | |
| REFERENCE | | | |
| AUTHORS | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. | | |
| TITLE | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer | | |
| JOURNAL | Genome Res. 10 (11), 1757-1771 (2000) | | |
| MEDLINE | 20530913 | | |
| PUBMED | 11076861 | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. | | |
| JOURNAL | Functional annotation of a full-length mouse cDNA collection | | |
| REFERENCE | Nature 409, 685-690 (2001) | | |
| AUTHORS | | | |
| TITLE | The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. | | |
| JOURNAL | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs | | |
| REFERENCE | Nature 420, 563-573 (2002) | | |
| AUTHORS | 6 (bases 1 to 3038) | | |
| TITLE | Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) | | |
| COMMENT | cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. | | |
| FEATURES | Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers | | |

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 8,34e-70 | Length: | 3038 |
| Score: | 636.50 | Matches: | 116 |
| Percent Similarity: | 75.37% | Conservative: | 37 |
| Best Local Similarity: | 57.14% | Mismatches: | 45 |
| Query Match: | 59.10% | Indels: | 5 |
| DB: | 11 | Gaps: | 2 |

US-10-029-020-14_COPY_2400_2600 (1-201) x AK037897 (1-3038)

| | | |
|----|------|--|
| QY | 1 | IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20 |
| DB | 1869 | ATTGGTTTTCATGGAGGACTATGATTTCTTACTAACTAGTGATCTAGGCGCAAGA 1928 |
| QY | 21 | AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGlyLeuTyrLysHisLeu 40 |
| DB | 1929 | GATTATGATGTTGTGTGCGCAGATGGACACGCTATCATCATCATATGGAACAGTTG 1988 |
| QY | 41 | SerSerSerAsnValMet-----ProPheAsnLeuTyrMetPheLysAsnAsnPro 58 |
| DB | 1989 | -----AACCTTCTTCTTAAACCATTCACCTCTACTCTTTGAGAATACTACCCA 2039 |
| QY | 59 | IleSerAsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuThr 78 |
| DB | 2040 | GTGGCAAAATTCAGAGTGTTCGAAGTATACCAAGATATCGGACCTGGCTGGAGCTC 2099 |
| QY | 79 | PheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMet 98 |
| DB | 2100 | TTTGGTTTCCAGTTACATAAGCTACTCTGGATTTCACCAACCAAGAACTGGAACAATG 2159 |
| QY | 99 | GluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLys 118 |
| DB | 2160 | GAATTAACTTATGAGCTTCTCAACTCGACAGACAAAACCAAGAGTGGGCTCTCTGAAAG 2219 |

119 SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGlu 138
 2220 ATGATCTCGGCATTCAGTGTGAGCTCCAGAAACAATTAAGGAATTTTCATTTCTGGAT 2279
 139 ArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLys 158
 2280 CAGCTTCCTATGACTCTCTCAGTACAAATGAGGGAGGTCTTGAAGGGAGAAACAGCCG 2339
 159 LysPheAlaSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGly 178
 2340 AGGTTTGTCTGCTCCCTCTGCTCTTGGCAAGGTATCAATTTGCCATCAAGGAGGC 2399
 179 ArgValThrThrAspIleSerValAlaLeuGlnAspGlyArgValAlaAlaLar 198
 2400 ATAGTACAGCTGATATTATAGGAGTAGCAAGTAAAGTAGCAGCGCTTGTCTGCCAAT 2459
 199 LeuAenHis 201
 2460 CTCACAAAT 2468

BU371366 929 bp mRNA linear EST 28-NOV-2002
 603598013F1 CSEQCHN73 Gallus gallus cDNA clone ChEST56814 5', mRNA
 sequence.
 BU371366
 BU371366.1 GI:25879367
 EST
 Gallus gallus (chicken)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 929)
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 2235534
 12445392
 Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1. 929
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Compton Line 151"
 /db_xref="taxon:9031"
 /clone="CHEST56814"
 /sex="Female"
 /tissue_type="not cerebrum or cerebellum"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN73"
 /note="Organ: brain; Vector: pBluescript II KS(+); Site 1:
 EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand reaction, double-stranded cDNA
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research
 (1996) 791, except that a significantly longer

FEATURES
 source
 RESULT 13
 BE819779/c
 LOCUS
 DEFINITION
 BE819779
 BE819779.1 GI:10252013
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 506)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

reannealing hybridization was used."

ORIGIN
 Alignment Scores: 5.65e-70 Length: 929
 Pred. NO.: 631.50 Matches: 111
 Score: 76.96% Conservativeness: 36
 Percent Similarity: 58.12% Mismatches: 43
 Best Local Similarity: 58.64% Indels: 1
 Query Match: 13 Gaps: 1
 DB:
 US-10-029-020-14_COPY_2400_2600 (1-201) x BU371366 (1-929)
 QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
 DB 84 ATTGGTTTTCATGGAGGCTCTATGATCTCTCACCATAATTTGGTTTCATCTGGGTGAGCGG 143
 QY 21 AspTyrAspValLeuAlaGlyArgThrPheSerProAspHisGluLeuTyrLysHisLeu 40
 DB 144 GATTATGATGTTATTGTGGTGGGACACACCAACCATCATATATGGAACACCTG 203
 QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProLysSer 60
 DB 204 AATGCTGTC--CCACACCATTCATCTCTACTTCTTGAATAAATACTACCAGTTGCC 260
 QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
 DB 261 AGGATCCAGATGTTGCTAAGTATACACAGACATTTGGAAGTTGGCTAGAGCTATTGGT 320
 QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 DB 321 TTCCAGTTCCACATGATCTACTCTGGTTTCCAAACACAGAAATCGAAGCTTTGGAGACA 380
 QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
 DB 381 ACATATGACTTCTACAGCTTCAACAAAAACCCAGAGTGGGATCCTCGAAAGACTATC 440
 QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 DB 441 CTGTGATTTCAGTGTGAGCTACAGAACCACTCCGAAACTTTATATATCTTGGATCACTT 500
 QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
 DB 501 CCAATGACCCCGAGGTATAGTGGCAAGTCTATGAGGAGTGAAGCAACCGAGGTTT 560
 QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
 DB 561 GCAGCTATTCTTTCAGTATTGGAAAGGCACTCAAAATTTGTATCAAGGATGGCATCGTC 620
 QY 181 ThrThrAspIleLeuSerValAlaAsnGluAsp 191
 DB 621 ACAGCGGACATTAATTGGTGTGGCTTAATGAGGAC 653

506 bp mRNA linear EST 21-SEP-2000
 BE819779
 MR3-BN0363-150800-006-c07 BN0363 Homo sapiens cDNA, mRNA sequence.
 BE819779
 BE819779.1 GI:10252013
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 506)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL
MEDLINE
PubMed
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=4t2-MR3-BN0363-150
800-006-c07&t3=2000-08-15&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 377.

FEATURES
source

1..506
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0363"
/note="Organ: breast_normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.: 4,54e-70 Length: 506
Score: 629.00 Matches: 122
Percent Similarity: 98.43% Conservative: 3
Best Local Similarity: 96.06% Mismatches: 2
Query Match: 58.40% Indels: 1
DB: 10 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x BE819779 (1-506)

QY 54 LysAsnAsnProIleSerAsnSerGlnAspIleLysCysPheMetThrAspValAsn 73
DB 386 CAAACACTACAAACCCCATCCAGCACTCCAGGACATCAAGTGCTTCATGACAGATGTTAAC 327
QY 74 SerTrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysPro 93
DB 326 AGCTGGCTGCTACCTTTGGATTCCAGCTATCCACACACAGATGAAGACGACGAG 267
QY 94 AspMetAspAlaMetGluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGlu 113
DB 266 GACATGGATGCCATGAACCTCTCAGAGCTATCCACACACAGATGAAGACGACGAG 207
QY 114 TrpAspAsnSerLysSerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAla 133
DB 206 TGGGCAACAGCAAGCTCTCTCTCGGGTGCAGTGAAGTACAGAAGCAGCTCAAGGCC 147
QY 134 PheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGln 153
DB 146 TTTGTACCTTAAACGGTTTGACAGCTCTATGGCTCCACCATCACCAGCTGCCAGCAG 87
QY 154 AlaProLysThrLysLysPheAlaSerSerGlySerValPheGlyLysValLysPhe 173
DB 86 GCTCAACAGACCAAGAGTTTGATCCACCGGCTCAGTCTTTGGCAG-GGAGTCAAGTTT 28
QY 174 AlaLeuLysAspGlyArgVal 180
DB 27 GCCTTGAAGGATGGCCGAGTA 7

RESULT 14
CB578819

LOCUS

CB578819 621 bp mRNA linear EST 03-APR-2003
AMGNNUC:NRHY7-00034-D3-A nrhy7 (10850) Rattus norvegicus cDNA clone
nrhy7-00034-d3 5', mRNA sequence.

ACCESSION

CB578819

VERSION

CB578819.1 GI:29522860

KEYWORDS

EST.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 621)

AUTHORS

Amgen EST Program.

TITLE

Amgen Rat EST Program

JOURNAL

Unpublished (2003)

COMMENT

Contact: Dan Fitzpatrick

FEATURES

source

1..621

Location/Qualifiers

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clones="nrhy7-00034-d3"

/clone_lib="nrhy7 (10850)"

/note="Vector: pSPORT1; Site_1: Sall; Site_2: NotI; W Rat

hypothalamus adult female Wistar rat avg. insert size 2.3

kb fraction 6 and 7"

ORIGIN

Alignment Scores:

Pred. No.: 1.4e-68 Length: 621
Score: 618.50 Matches: 114
Percent Similarity: 74.87% Conservative: 35
Best Local Similarity: 57.29% Mismatches: 45
Query Match: 57.43% Indels: 5
DB: 14 Gaps: 2

US-10-029-020-14_COPY_2400_2600 (1-201) x CB578819 (1-621)

QY 5 GlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArgAspTyrAspVal 24
DB 1 GGAGGACTCTATGATTTCTTACTAACTAGTGCATCTAGGCAAGGGATTATGATTT 60
QY 25 LeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeuSerSerAsn 44
DB 61 GTTGCTGGCAGATGGCAACGCCCTAATCATCATATGGAACAGTTG-----AAC 111
QY 45 ValMet-----ProPheAsnLeuTyrMetPheLysAsnAsnProIleSerAsnSer 62
DB 112 CTCTCTCTAAACCACTCAACCTCTACTCTCTTGAGATAACTATCCGTTGGCAAAAT 171
QY 63 GlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGlyPheGln 82
DB 172 CAAGACGTTGCAAGTATACACACAGATGGCAGTTGGCTGGAGCTCTTTGGTTCCAA 231
QY 83 LeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluProSerTyr 102
DB 232 TTACATACGCTGCTACCTGGATTCCCAACACAGAACTAGAAAAATGGAATTAAT 291
QY 103 GluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIleLeuGly 122
DB 292 GAGCTTCTTTCAGCTTCACACAAAACCCAGAGTGGGATCTCCGGCAAGATGATCTGGGC 351
QY 123 ValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPheAspGln 142
DB 352 ATTCAGTGTGAGCTCCAGAAACAGCTCAGGAATTTTCATTTCTTGATCAGCTTCTATG 411
QY 143 LeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPheAlaSer 162
DB 412 ACTCCTCAGTACATAGAGGGCGGTGGCTTGAAGGAGGGAACACAGCAAGGTTTGCTGCT 471

1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrIleValHisMetGlyArg 20
 650 ATTGGATTTCATGGTGGCTGTATGACCCACTACCAAAATTAATCCACTTTGGGAAAGA 591
 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGlyLeuTyrLysHisLeu 40
 590 GATTATGACATTTTGGCAGGACGGTGGACAACTGACATAGAAATCTGGAAGAAT 531
 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProLysSer 60
 530 ---GGGAGGAGCCAGCTCCTTTTAACTGTGTACATGTTTAGGAATAACCAACCTGCAAGC 474
 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrLeuLeuThrPheGly 80
 473 AATATCCATGACGTGAAAGATTACATCACAGATGTTAAACAGCTGGCTGGTACATTTGGT 414
 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 413 TTCCATCTGCAAAATGCTTATTCCTGGATTCCTGTCCCAAAATTTGATTTAAACAGAACT 354
 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
 353 TTTTACGAACTTGTG-----AAGAGTCAGCAGTGGGATCATATACCGCCATC 306
 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 305 TTTCGAGTCCAGCAGCAAGTGGCGGCGAGGCGCTTCTGTCTGGTGGGGAAGATG 246
 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
 245 GCCGAGGTGCAG-----GTGAGCCGCGCGCGCGCGCGCGCGCTCTGG 198
 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
 197 CTGTGTTTCCGACCGTCAAGTCGTGATCGGCAAGGGCGTTCATGTGCGCGCTGACGCCAG 138
 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAla 197
 137 GGCGCGGTGCGACACCAACGTCGTCAACATCGCCACAGAGAGCTGCATCAAGTGGCGGCC 78
 198 IleLeuAsnHis 201
 77 GTGCTCAACAAC 66

Search completed: August 14, 2004, 18:04:53
 Job time : 2373.29 secs

QY 163 SerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgValThrThr 182
 Db 472 GTCCCTCTGCTTTGGCAAGGTATAAAATTTGCCATCAAGGAGGCATAGTACAGCT 531
 QY 183 AspIleLeuSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsnHis 201
 Db 532 GATATTAGGAGTAGCAATGAAGATAGCAGGCGCTCTGTGCTCCATTTCTCACAAT 588

RESULT 15
 CA777388 776 bp mRNA linear EST 03-DEC-2002
 LOCUS ip19c06.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6217715 3'
 DEFINITION similar to TR:Q9WTS6 Q9WTS6 TEN-M3.1; mRNA sequence.

ACCESSION CA777388
 VERSION CA777388.1 GI:26015263
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 776)
 AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scedarce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Teagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
 Endocrine Pancreas Consortium

TITLE Unpublished (2000)
 JOURNAL Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 COMMENT Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@im.wustl.edu)
 Seq primer: -400P from Gibco
 High quality sequence stop: 439.
 Location/Qualifiers
 1..776
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6217715"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_lib="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
 NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
 Size-selected on agarose gel. Average insert size ~1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."

FEATURES

source

ORIGIN

Alignment Scores:
 Align. No.: 1,12e-61 Length: 776
 Score: 566.00 Matches: 107
 Percent Similarity: 71.08% Conservative: 38
 Best Local Similarity: 52.45% Mismatches: 47
 Query Match: 52.55% Indels: 12
 DB: 14 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x CA777388 (1-776)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 340.95 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_2400_2600
Perfect score: 1077
Sequence: 1 IGVHGGYDLTKLVHGR.....TDIISVANEDGRRVAAILNH 201

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N_Geneseq_29Jan04 -QPM=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CGN 1.1.1868 @runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*

1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001as:.*
5: Geneseq2001bs:.*
6: Geneseq2002s:.*
7: Geneseq2003as:.*
8: Geneseq2003bs:.*
9: Geneseq2003cs:.*
10: Geneseq2004as:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----------|--------------------|
| 1 | 1077 | 100.0 | 8354 | ABS52100 | ABS52100 Human TEN |
| 2 | 1077 | 100.0 | 8645 | ABS78652 | ABS78652 Human CDN |
| 3 | 1063 | 98.7 | 8438 | ABN85378 | ABN85378 Human NOV |
| 4 | 642.5 | 59.7 | 13202 | AAK51828 | AAK51828 Human pol |
| 5 | 630.5 | 58.5 | 12879 | ABK92230 | ABK92230 Prostate |
| 6 | 566 | 52.6 | 2157 | AAH14096 | AAH14096 Human CDN |
| 7 | 566 | 52.6 | 3270 | AAH14671 | AAH14671 Human CDN |
| 8 | 566 | 52.6 | 3614 | AAH14183 | AAH14183 Human CDN |

| | | | | | | |
|----|-------|------|--------|---|----------|--------------------|
| 9 | 566 | 52.6 | 8473 | 6 | ABQ82345 | Abq82345 Human NOV |
| 10 | 566 | 52.6 | 8487 | 6 | ABQ82346 | Abq82346 Human NOV |
| 11 | 566 | 52.6 | 8645 | 6 | ABQ82344 | Abq82344 Human NOV |
| 12 | 566 | 52.6 | 8675 | 6 | ABQ82343 | Abq82343 Human NOV |
| 13 | 556 | 51.6 | 791 | 6 | ABK34316 | Abk34316 Human CDN |
| 14 | 529.5 | 49.2 | 9058 | 7 | ACC72051 | Acc72051 BCU0205A |
| 15 | 529.5 | 49.2 | 9695 | 7 | ACC72052 | Acc72052 BCU0205B |
| 16 | 529.5 | 49.2 | 9729 | 5 | AAI14089 | Aai14089 Human FCT |
| 17 | 529.5 | 49.2 | 9729 | 9 | ADB32028 | Adb32028 Human FCT |
| 18 | 529.5 | 49.2 | 9826 | 5 | AAI14085 | Aai14085 Human FCT |
| 19 | 529.5 | 49.2 | 9826 | 9 | ADB32023 | Adb32023 Human FCT |
| 20 | 497.5 | 46.2 | 3312 | 4 | AAK52812 | Aak52812 Human pol |
| 21 | 200.5 | 18.6 | 10242 | 4 | ABL29075 | Abi29075 Drosophil |
| 22 | 200.5 | 18.6 | 17131 | 4 | ABL29074 | Abi29074 Drosophil |
| 23 | 194 | 18.0 | 431 | 6 | ABL65724 | Abi65724 Lung canc |
| 24 | 188.5 | 17.5 | 1973 | 4 | AAI14695 | Aai14695 Probe #46 |
| 25 | 188.5 | 17.5 | 1973 | 4 | ABA56427 | Aba56427 Human foe |
| 26 | 188.5 | 17.5 | 1973 | 4 | AAI36063 | Aai36063 Probe #47 |
| 27 | 188.5 | 17.5 | 1973 | 4 | ABA45900 | Aba45900 Human bre |
| 28 | 188.5 | 17.5 | 1973 | 4 | ABA26060 | Aba26060 Probe #45 |
| 29 | 188.5 | 17.5 | 1973 | 4 | AAK30099 | Aak30099 Human bon |
| 30 | 188.5 | 17.5 | 1973 | 4 | AAK04592 | Aak04592 Human bra |
| 31 | 188.5 | 17.5 | 1973 | 4 | ABS29749 | Abs29749 Human liv |
| 32 | 188.5 | 17.5 | 1973 | 5 | AAI04501 | Aai04501 Probe #44 |
| 33 | 188.5 | 17.5 | 1973 | 6 | ABS04679 | Abs04679 Human gen |
| 34 | 174 | 16.2 | 728 | 5 | AAI04501 | Aai04501 Probe #44 |
| 35 | 164.5 | 15.3 | 3910 | 4 | AAI04501 | Aai04501 Probe #44 |
| 36 | 164.5 | 15.3 | 6318 | 4 | ABL04848 | Abi04848 Drosophil |
| 37 | 131.5 | 12.2 | 399 | 6 | AAI04848 | Abi04848 Drosophil |
| 38 | 128 | 11.9 | 464 | 8 | AAI04848 | Abi04848 Drosophil |
| 39 | 87.5 | 8.1 | 2836 | 7 | ABX70662 | Abx70662 Human CDN |
| 40 | 86 | 8.0 | 1630 | 4 | ABL03289 | Abi03289 Drosophil |
| 41 | 84.5 | 7.8 | 1222 | 4 | AAI03289 | Abi03289 Drosophil |
| 42 | 84.5 | 7.8 | 1527 | 7 | ABX06099 | Abx06099 S. pneumo |
| 43 | 84.5 | 7.8 | 20986 | 2 | AAV52187 | Aav52187 Streptoco |
| 44 | 84.5 | 7.8 | 110000 | 7 | ABS56454 | Abss56454 03 |
| 45 | 84 | 7.8 | 2419 | 3 | AAZ91917 | Aaz91917 Murine ma |

ALIGNMENTS

RESULT 1
ABS52100
ID ABS52100 standard; DNA; 8354 BP.
AC ABS52100;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human TEN-M4-like gene.
XX

Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
cell signal processing; metabolic pathway modulation; metabolic disorder;
obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
memory defect; infertility; congenital heart defect; hair growth;
pigmentation disorder; endocrine disorder; respiratory disease; health;
gastro-intestinal disease; reproductive; neurological disease;
bone marrow transplantation; endocrine disease; allergy; inflammation;
nephrological disorder; urinary system disorder; age-related disorder;
neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
adipocyte complement-related C1q tumour necrosis factor; out at first;
beta adrenergic receptor kinase; BphA6/ehk-2; glucose transporter;
type Ia membrane sushi-containing domain; butyrophilin;
single nucleotide polymorphism.

Homo sapiens.

Key Location/Qualifiers
variation replace(117,G)
/*tag= a

variation /standard_name="Single nucleotide polymorphism (SNP)"
 replace(225,C)
 /*tag= b
 /standard_name="Single nucleotide polymorphism (SNP)"
 replace(260,A)
 /*tag= c
 /standard_name="Single nucleotide polymorphism (SNP)"

WO200257453-A2.

25-JUL-2002.

19-DEC-2001; 2001WO-US050331.

19-DEC-2000; 2000US-0256704P.

20-DEC-2000; 2000US-0257314P.

02-MAY-2001; 2001US-0288153P.

29-MAY-2001; 2001US-0294075P.

24-JUL-2001; 2001US-0307506P.

10-AUG-2001; 2001US-0311590P.

10-AUG-2001; 2001US-0311613P.

29-AUG-2001; 2001US-0315617P.

14-SEP-2001; 2001US-0322358P.

(CURA-) CURAGEN CORP.

Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;

Stone DJ, Anderson D, Shmukets RA, Burgess CE, Zerhusen BD, Liu X;

Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;

WPI; 2002-590744/63.

P-PSDB; ABG70388.

Novel isolated NOVX polypeptide useful for treating cardiomyopathy, atherosclerosis, metabolic disorders, diabetes, obesity, infectious disease, anorexia, neurodegenerative disorders, Alzheimer's disease or cancer.

Claim 8; Page 50-52; 318pp; English.

The present invention relates to new NOVX polypeptides. The invention is useful for treating or preventing a NOVX-associated disorder such as cardiomyopathy or atherosclerosis, where the disorder is related to cell signal processing and metabolic pathway modulation in a subject. Preferably human. The invention is also useful for treating metabolic disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disease), immune disorders, haematopoietic disorders and various cancers. The molecules of the invention are also useful for treating or preventing cirrhosis, pancreatitis, learning and memory defects, infertility, congenital heart defects, acne, hair growth, pigmentation disorders, endocrine disorders, respiratory diseases, gastro-intestinal diseases, reproductive health, neurological diseases, bone marrow transplantation, endocrine diseases, allergy and inflammation, nephrological disorders, urinary system disorders, neuropsychiatric disorders and age-related disorders. The present nucleic acid sequence represents a NOVX gene. This sequence encodes a NOVX protein of the invention

Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 6.8e-122 Length: 8354
 Score: 1077.00 Matches: 201
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x ABS52100 (1-8354)

1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20

7232 ATAGCTTACCATGCTGCTATGATCCACTCAAGCTTGTCCACATGGCCGGCGA 7291

QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
 Db 7292 GATTATGATGCTGCTGGCGGACGCTGCTAGCCAGCACACAGAGCTGTGGAAGCACTT 7351
 QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
 Db 7352 AGTAGCAGCAACGCTCATGCTCTTTTAATCTCTATATGTTCAAAAACAACACCCATCAGC 7411
 QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrLeuLeuThrPheGly 80
 Db 7412 AACTCCAGGACATCAAGTGTCTTATGACAGATGTTTACAGCTGGCTGCTCACCTTTGGA 7471
 QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 Db 7472 TTCCAGCTACACACGCTGATCCCTGGTTATCCCAAAACAGACATGATGATGGAACCC 7531
 QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
 Db 7532 TCCTACGAGCTCATCCACACACAGATGAAACGCGAGGAGTGGGACACAGCAAGTCTATC 7591
 QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 Db 7592 CTCGGGGTACAGTGTGAAGTACAGACAGCTCAAGGCTTTGTACACCTTAGAACGGTTT 7651
 QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
 Db 7652 GACCAGCTCTATGGCTCCACAAATCACCAGCTGCCAGCAGGCTCCAAAGACCAAGAAGTTT 7711
 QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
 Db 7712 GCATCCAGCGGCTCAGTCTTTTGGCAAGGGGGTCAAGTTTGCCTTGAAGATGCCGAGTG 7771
 QY 181 ThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAlaLeuLeuAsn 200
 Db 7772 ACCACAGACATCATGATGTCCTCAATGAGGATGGCGAGGTTGCTGCCATCTTTGAAC 7831
 QY 201 His 201
 Db 7832 CAT 7834

RESULT 2

ABS78652

ID ABS78652 standard; cDNA; 8645 BP.

AC ABS78652;

XX ABS78652;

DT 16-DEC-2002 (first entry)

XX Human cDNA encoding CGDD10, INCVTE 7488573CB1.

Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
 cell proliferative disorder; arteriosclerosis; atherosclerosis;
 cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
 polycythaemia vera; primary thrombocytopaenia; developmental disorder;
 renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
 neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
 reproductive disorder; infertility; autoimmune disorder; gout; allergy;
 inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
 autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
 diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
 multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
 rheumatoid arthritis.

XX Homo sapiens.

OS WO200272830-A2.

XX 19-SEP-2002.

XX 08-FEB-2002; 2002WO-US003715.

XX 09-FEB-2001; 2001US-0268111P.

XX PD

XX PF

XX XX

XX PR

PR 23-FEB-2001; 2001US-0271175P.
 PR 08-MAR-2001; 2001US-0274503P.
 PR 09-MAR-2001; 2001US-0274552P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
 PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
 PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
 XX
 DR WPI; 2002-723356/78.
 DR P-PSDB; ABG97359.
 XX
 XX New human proteins associated with cell growth, differentiation and
 PT death, useful for diagnosing, treating or preventing autoimmune or
 PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
 PT atherosclerosis or hepatitis.
 XX
 PS Claim 5; Page 175-178; 181pp; English.
 XX
 CC The invention relates to an isolated polypeptide comprising CGDD1-12
 CC (cell growth, differentiation and death), a naturally occurring amino
 CC acid sequence at least 90% identical to CGDD, a biologically active
 CC fragment or an immunogenic fragment. Also included are the
 CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
 CC comprising a promoter sequence operably linked to the CGDD
 CC polynucleotides, a cell transformed with the recombinant polynucleotide,
 CC a transgenic organism comprising the recombinant polynucleotide, an anti-
 CC CGDD antibody, screening for compounds which bind to/modulate or are
 CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
 CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
 CC agonists and antagonists are useful for diagnosing, treating or
 CC preventing disorders associated with aberrant expression of CGDD,
 CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
 CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
 CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
 CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
 CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
 CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
 CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
 CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
 CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
 CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
 CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
 CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
 CC infections. They are also useful in the assessment of the effects of
 CC exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of proteins associated with CGDD. The present sequence encodes
 CC a CGDD protein
 XX
 SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;
 XX
 Alignment Scores:
 Pred. No.: 7,14e-122 Length: 8645
 Score: 1077.00 Matches: 201
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-020-14_COPY_2400_2600 (1-201) x ABS78652 (1-8645)
 QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
 Db ATAGGCTACCATGGTGGCTCTATGATCCACTCACCAGCTTGTCACATGGCGCGGA 7340
 QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
 Db GATTATGATGTGCGCGGACGCTGGATAGCCAGACCCAGAGCTGTGGAGCACCTT 7400
 QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
 Db AGTAGCAGCAAGCTCATGCCCTTTTAACTCTATATATGTTCAAAACCAACACCCCATCAGC 7460

QY 61 AnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuThrPheGly 80
 Db AACTCCAGGACATCAAGTGTTCATGACAGATGTTACAGCTGGCTGCTACCTTTGA 7520
 QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 Db TTCCAGCTACACAAACGCTGATCCCTGGTTATCCCAACACAGCATGGATGCCATGAAACC 7580
 QY 101 SerTyrGluLeuLeuHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
 Db TCCTACGAGCTCATCCACACACATGAACAGGAGGTGGGCAACAGCAAGTCTATC 7640
 QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 Db CTCGGGTACAGTGTGAAGTACAGAGCAGCTCAAGGCTTTGTACCTTGAACGGTTT 7700
 QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
 Db GACCAGCTCTATGGCTCCCAATCACCAGCTGCCAGAGGCTCCAAAGACCAAGAGTTT 7760
 QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
 Db GCATCCAGCGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTG 7820
 QY 181 ThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
 Db ACCACAGACATCATCAGTGTGGCAATGAGGATGGCGGAGGGTTGCTGCCATCTTGAAAC 7880
 QY 201 His 201
 Db CAT 7883
 XX
 RESULT 3
 ABN85378
 ID ABN85378 standard; DNA; 8438 BP.
 XX
 AC ABN85378;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human NOV1, TEN-M4 like protein, coding sequence.
 XX
 KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
 KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
 KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
 KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
 KW Gene Therapy; NOV; cancer; heart disease; inflammation;
 KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
 KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
 KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
 KW TEN-M4 like protein; chromosome 11; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 4..8395
 FT /*tag= a
 FT /trans except= (pos: 1138..1147,aa:Met)
 FT /product= "NOV1 protein"
 XX
 WO200255704-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2002; 2002WO-US000554.
 XX
 PR 09-JAN-2001; 2001US-0260417P.
 PR 10-JAN-2001; 2001US-0260831P.
 PR 28-FEB-2001; 2001US-0272338P.
 PR 09-MAR-2001; 2001US-0274876P.
 PR 18-APR-2001; 2001US-0284704P.
 XX

(CURA-) CURAGEN CORP.

PA Padigar M, Li L, Zerhusen BD, Casman SU, Shenoy S, Spytek KA;
 XX Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;
 PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
 PI Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
 PI Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
 PI Gunther E, Stone DJ;
 DR WPI: 2002-590674/63.
 DR P-PSDB: ABB98401.

NOVX polypeptides and encoding polynucleotides, useful for preventing or treating NOVX-associated disorders e.g. cancer, inflammation, or Alzheimer's disease, and in chromosome mapping, tissue typing or pharmacogenomics.

Claim 9; Page 8-9; 358pp; English.

The present sequence is a coding sequence for a NOV protein. The NOV proteins and coding sequences are useful for treating or preventing NOV-associated disorders or in the manufacture of a medicament for treating the disorders, such as cancer, heart disease, inflammation, autoimmune disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma, IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy, and other wasting disorders associated with chronic diseases. NOV1 is a TEN-M4 like protein and the NOV1 gene is localised to chromosome 11

Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 3,74e-120 | Length: | 8438 |
|------------------------|-----------|---------------|------|
| Score: | 1063.00 | Matches: | 201 |
| Percent Similarity: | 96.17% | Conservative: | 0 |
| Best Local Similarity: | 96.17% | Mismatches: | 0 |
| Query Match: | 98.70% | Indels: | 8 |
| DB: | 6 | Gaps: | 1 |

US-10-029-020-14_COPY_2400_2600 (1-201) x ABN85378 (1-8438)

QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
 DB 7259 ATAGGCTACCATGGTGGCTCTATGATCCACTCACCACTGTCTCATGGCGCGCA 7318

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
 DB 7319 GATTATGATGTCTGCGCGGACGCTGACTAGCCAGACCCAGAGCTGTGGAAGCACCTT 7378

QY 41 SerSerSerAsnValMetProProHeAnLeuTyrMetPhelysAsnAsnProLysSer 60
 DB 7379 AGTAGCAGCAAGCTCATGCTCTTTAATCTCTATATCTTCAAAACCAACACCCATCAGC 7438

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
 DB 7439 AATCCACAGACATCAAGTCTTCATGACAGATGTTAAACAGCTGGCTGCCTTTGGA 7498

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 DB 7499 TTCACGCTACACAACTGATCCCTGGTGTATCCAAACACAGACATGGATGCAATGGAACCC 7558

QY 101 SerTyrGluLeuLeuHisThrGlnMetLysThrGlnGluTrpAspAsnSerLys----- 118
 DB 7559 TCCTACGAGCTCATCACACACAGATGAAACGAGGAGTGGGACACACACAGGTAAATT 7618

QY 119 -----SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLys 132
 DB 7619 CTGACACAGCTGCCAGTCTATCTCTCGGGTACAGTGTGAAGTACAGAGCAGCTCAAG 7678

QY 133 AlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGln 152
 DB 7679 GCCTTTGCTACCTTAGAACGGTTTGAACAGCTCTATGGCTCCCAATCACCAGCTGCCAG 7738

QY 153 GlnAlaProLysThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyValLys 172
 DB 7739 CAGGCTCCAAAGACCAAGAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGCTCAAG 7798

QY 173 PheAlaLeuLysAspGlyArgValThrThrAspIleIleSerValAlaAsnGluAspGly 192
 DB 7799 TTTGCTTGAAGGATGCGCGAGTGACCAACAGACATCATCATGTGTGGCCATGAGATGGG 7858

QY 193 ArgArgValAlaAlaIleLeuAsnHis 201
 DB 7859 CGAAGGGTGTGCTGCACTTGAACCAT 7885

RESULT 4
 AAK51828
 ID AAK51828 standard; cDNA; 13202 BP.
 XX AAK51828;
 AC AAK51828;
 XX 06-NOV-2001 (first entry)
 DT Human polynucleotide SEQ ID NO 373.
 DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX Homo sapiens.
 OS WO200157190-A2.
 PN 09-AUG-2001.
 PD 05-FEB-2001; 2001WO-US004098.
 PP 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
 XX WPI: 2001-476283/51.
 DR P-PSDB: AAM78695.
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.
 PT Claim 1; Page 1414-1426; 6221pp; English.
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and/or activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication
 CC Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;
 XX

| XX | PF | 12-OCT-2001; 2001WO-US032045. |
|--|---|-------------------------------|
| XX | 13-OCT-2000; 2000US-00687576. | |
| PR | 08-DEC-2000; 2000US-00733288. | |
| PR | 08-DEC-2000; 2000US-00733288. | |
| PR | 24-JAN-2001; 2001US-02639572. | |
| PR | 16-MAR-2001; 2001US-02639572. | |
| PR | 16-MAR-2001; 2001US-02639572. | |
| PR | 06-APR-2001; 2001US-02639572. | |
| PR | 24-APR-2001; 2001US-02639572. | |
| PR | 30-APR-2001; 2001US-02639572. | |
| PR | 04-MAY-2001; 2001US-02639572. | |
| XX | (EOSB-) EOS BIOTECHNOLOGY INC. | |
| XX | Gish KC, Mack DH, Wilson KE, Afar D, Heverzi P; | |
| XX | WPI; 2002-471335/50. | |
| XX | P-PSDB; ABG61913. | |
| PT | Detecting a prostate cancer-associated transcript in a cell in a patient, | |
| PT | useful for diagnosing prostate cancer (PC) or screening modulators of PC, | |
| PT | by determining if prostate cancer-associated genes are expressed in a | |
| PT | prostate tissue. | |
| XX | Claim 22; Page 394-397; 436pp; English. | |
| XX | The present invention relates to methods of detecting a prostate cancer- | |
| CC | associated transcript in a cell from a patient. The method comprises | |
| CC | contacting a biological sample from the patient with prostate cancer- | |
| CC | associated polynucleotides (designated PC genes) that selectively | |
| CC | hybridize to a sequence that is at least 80% identical to them. The | |
| CC | prostate cancer-associated polynucleotide sequences are differentially | |
| CC | expressed in prostate tumour tissue or in prostate cancer and are derived | |
| CC | from the tissues of various organisms such as humans or other mammals | |
| CC | (e.g. mice, sheep and dogs). The methods of the invention are useful for | |
| CC | diagnosing and treating prostate cancer in mammals. The prostate cancer- | |
| CC | associated genes are useful for diagnosing or treating prostate cancer, | |
| CC | as well as for identifying modulators of prostate cancer or agents that | |
| CC | inhibit prostate cancer. The nucleic acid sequences are particularly | |
| CC | useful in gene therapy, as a vaccine or in antisense applications. | |
| CC | ABK92115-ABK92263 represent prostate cancer-associated polynucleotide | |
| CC | sequences | |
| XX | Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other; | |
| SQ | Alignment Scores: | |
| | Pred. No.: 2,62e-66 Length: 12879 | |
| | Score: 630.50 Matches: 116 | |
| | Percent Similarity: 74.88% Conservative: 36 | |
| | Best Local Similarity: 57.14% Mismatches: 46 | |
| | Query Match: 58.54% Indels: 5 | |
| | DB: 6 Gaps: 2 | |
| US-10-029-020-14_COPY_2400_2600 (1-201) x ABK92230 (1-12879) | | |
| QY | 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgG 20 | |
| DB | 7133 ATTGGTTTCATGGAGGACTCTATGATTTCTTAAATTTAGTGACCTGGGCAAGG 7192 | |
| QY | 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40 | |
| DB | 7193 GATTATGATGTTGTTGCTGGCAGATGGACACCGCCCTATCATCATATGGAACAGTTG 7252 | |
| QY | 41 SerSerSerAsnValMet-----ProPheAsnLeuTyrMetPheLysAsnAsnPro 58 | |
| DB | 7253 -----AACTCTCTCTAAACCATTCACCTCTACTCTCTTTGAAATAATACCTCCA 7303 | |
| QY | 59 IleSerAsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrLeuLeuThr 78 | |
| DB | 7304 GTTGGCAAAATTCAGATGTTGCAAGATGTTACCAACAGATACCAACAGATGTTGGAGCTA 7363 | |
| XX | 18-APR-2002. | |

79 PheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMet 98
 7364 TTGGTTTCCAAATACACAATGACTACTGATTTCCCAACCTGAATAGAAAATTTA 7423
 99 GluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLys 118
 7424 GAATTAATCTACGAGCTCTACGCGCTTCAGACAAAACCTCAAGAGTGGATCTCTGGAAG 7483
 119 SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGlu 138
 7484 ACTATCTGGGCACTCAGTGTGAATCAGAAACAGCTCAGGAATTTCAATTCCTTGAC 7543
 139 ArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLys 158
 7544 CAACACTATGACTCTCCCGGTACAAATGATGAGCGGTGCTTGAAGGAGGAGCAACCA 7603
 159 LysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGly 178
 7604 AGGTTTGTCTGTCTCCCTCTCTGTTTGGGAAGGTATAAATTTGCCATCAAGGATGSC 7663
 179 ArgValThrThrAspIleSerValAlaAsnGluAspGlyArgValAlaAlaIle 198
 7664 ATAGTAACAGCTGATATTATAGGAGTACCAATGAAGATAGCAGCGGCTTGCTGCCATT 7723
 199 LeuAsnHis 201
 7724 CTCATAAT 7732

RESULT 6
 AAH14096
 ID AAH14096 standard; cDNA; 2157 BP.
 AC AAH14096;
 XX
 XX 26-JUN-2001 (first entry)
 DE Human cDNA sequence SEQ ID NO:11265.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX Homo sapiens.
 XX BP1074617-A2.
 XX 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-00116126.
 XX 29-JUL-1999; 99JP-00248036.
 XX 27-AUG-1999; 99JP-00300253.
 XX 11-JAN-2000; 2000JP-00118776.
 XX 02-MAY-2000; 2000JP-00183767.
 XX 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX Claim 8; SEQ ID NO 11265; 2537pp + Sequence Listing; English.
 XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence can be used in antisense therapy and in specification. The primer sets can be used for synthesizing polynucleotides, gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialized methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
 XX
 SQ Sequence 2157 BP; 544 A; 560 C; 591 G; 462 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.9e-59 Length: 2157
 Score: 566.00 Matches: 107
 Percent Similarity: 71.08% Conservative: 38
 Best Local Similarity: 52.45% Mismatches: 47
 Query Match: 52.55% Indels: 12
 DB: 4 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x AAH14096 (1-2157)
 QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
 Db 818 ATTGGATTTCATGTTGGGCTGTATGACCCACTCACCAATTAATCCACTTTGGAGAA 877
 QY 21 AspTyrAspValLeuAlaGlyArgThrSerProAspHisGluLeuThrLysHisLeu 40
 Db 878 GATTATGACATTTTGGGAGCGGTGGAGACACCTGACATAGAAATCTGGAAAGAA 937
 QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProLys 60
 Db 938 ---GGGAGGACCCAGCTCTTTTAACTTGTATGTTTGGAAATAACAACCTGCAAGC 994
 QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
 Db 995 AAAATCCATGACGTGAAGATTACATACAGATGTTTAAACAGCTGGTGGTGCATTTGGT 1054
 QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 Db 1055 TTCCATCTGCACAAATCTATTCCTGGATTCCTGTTCCCAAAATTTGATTTAACAGAACCT 1114
 QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
 Db 1115 TCTTACGAACTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 1162
 QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 Db 1163 TTCCGAGTCAGCAGCAAGTGGCGCGCCAGCCCTTCTCTGCTGGGGAAGATG 1222
 QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
 Db 1223 GCCAGGTGCAG-----GTGAGCGCGCGCGCGCGCGCGCGCTCTGG 1270
 QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
 Db 1271 CTGTGGTTCCCGCGGTCAAGTCGCTGATCGGAAGGCGCTCATGTCTGGCGCTCAGCCAG 1330
 QY 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAla 197
 Db 1331 GGCGCGCTGAGCAACGCTGCTCAATCGCAACGAGGACTGCATCAGTCAAGTGGCGCC 1390
 QY 198 IleLeuAsnHis 201
 Db 1391 GTGCTCAACAAC 1402

```
RESULT 7
AAH14671
ID AAH14671 standard; cDNA; 3270 BP.
XX
AC AAH14671;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12354.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
FN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 12354; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primer sets are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 3270 BP; 867 A; 821 C; 852 G; 730 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,49e-59 Length: 3270
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 4 Gaps: 4
```

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPT; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID NO 11430; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the polynucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX Sequence 3614 BP; 967 A; 898 C; 928 G; 821 T; 0 U; 0 Other;

| Alignment Scores: | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 4.03e-59 | Length: | 3614 |
| Score: | 566.00 | Matches: | 107 |
| Percent Similarity: | 71.08% | Conservative: | 38 |
| Best Local Similarity: | 52.45% | Mismatches: | 47 |
| Query Match: | 52.55% | Indels: | 12 |
| DB: | 4 | Gaps: | 4 |

US-10-029-020-14_COPY_2400_2600 (1-201) x AAH14183 (1-3614)

| | | | |
|----|------|---|------|
| Qy | 1 | IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg | 20 |
| Db | 2192 | ATTGGATTTTCATGTGTGGCCGTGTATGACCCACTTACCAATTAATCCATTTGAGAAAGA | 2251 |
| Qy | 21 | AspTyrAspValLeuAlaGlyArgTyrSerProAspHisGluLeuTyrLysHisLeu | 40 |
| Db | 2252 | GATTATGACATTTTGGCAGGACGGTGGACAAACCTGACATAGAAATCTGGAAAGAAT | 2311 |
| Qy | 41 | SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer | 60 |
| Db | 2312 | ---GGGAAGGACCCAGCTCCCTTTAACTGTGTACATGTTTAGGAATTAACAACCTGCAAGC | 2368 |
| Qy | 61 | AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly | 80 |
| Db | 2369 | AAATCCATGACGTGAAGACATTACACAGATGTTAACAGCTGGCTGTGTGACATTTGGT | 2428 |
| Qy | 81 | PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro | 100 |
| Db | 2429 | TTCCATCTGCACATGTCTATTCCTGGATTCCCTGTTCCCAAAATTTGATTAAACAGAACCT | 2488 |
| Qy | 101 | SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle | 120 |
| Db | 2489 | TCTTACGAACTTGTG-----AAGATGACGAGTGGGATGATATACCGCCCATC | 2536 |

PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
 PI Malyankar U, Shinkets R, Tchernev VT, Spaderna SK, Gorman L;
 PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
 PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
 XX Gunther E, Smithson G, Millet I, Macdougall JR;
 DR MPI: 2002-732706/79.
 DR P-PSDB; ABP53588.

XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
 PT associated disorders, such as cancers, neurological disorders, disorders
 PT of vesicular transport, gastrointestinal disorders, and autoimmune
 PT diseases.

XX Claim 8; Page 119-121; 444pp; English.

XX The present invention describes novel human proteins designated NOVX,
 CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
 CC cytotatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,
 CC antidabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
 CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
 CC protozoacide and antihelminthic activities, and can be used in gene
 CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease selected from NOVX-associated disorder, such as cancers
 CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
 CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
 CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
 CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
 CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
 CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
 CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
 CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
 CC and protozoal infections. The NOVX proteins can be used as immunogens to
 CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
 CC used in chromosome mapping, identifying individuals from minute
 CC biological samples (tissue typing), and in forensic identification of a
 CC biological sample. The present sequence encodes human NOV15c, which is
 CC located on chromosome 4

XX SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 1.39e-58 Length: 8473
 Score: 566.00 Matches: 107
 Percent Similarity: 71.08% Conservative: 38
 Best Local Similarity: 52.45% Mismatches: 47
 Query Match: 52.55% Indels: 12
 DB: 6 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x ABQ82345 (1-8473)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
 Db 7050 ATTGATTTCATGGTGGCTGATGACCCACCTACCAATTAATCCACTTTGGAGAAAGA 7109
 QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGlyLeuThrLysHisLeu 40
 Db 7110 GATTATGACATTTTGGCAGGAGCGTGGACACACCTGCATAGAAATCTGGAAGAAAT 7169
 QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
 Db 7170 ---GGGAAGGAGCCAGCTCTTTAACTGTACATGTTTAGGAATAACACCTCAGC 7226
 QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrLeuThrPheGly 80
 Db 7227 AAATCCATGACGTGAAGAGATTACATACACATGTTAAACAGCTGCTGCATTTGGT 7286
 QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 Db 7287 TTCCATCTGCACATGCTATTCTCTGGATTCCTGTTCCCAAATTTGATTAAACAGAACT 7346
 QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120

Db 7347 TCTTACGAACTTGTG-----AAGAGTCACAGTGGGATGATATACGCCCATC 7394
 QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 Db 7395 TTCGGAGTCCACAGCAAGTGGCGGAGGCAAGGCCCTCTGCTGCTGGGGAAGATG 7454
 QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
 Db 7455 GCCGAGGTGCAG-----GTGAGCCGCCGCCGCCGCCGCCGCCGAGTCTCTGG 7502
 QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
 Db 7503 CTGTGTTTCGCGACGGTCAAGTCGTGATCGCGAAGGGCGTCTGCTGCGCTCAGCCAG 7562
 QY 178 GlyArgValThrAspIleSerValAlaGlnGluAspGlyArgValAlaAla 197
 Db 7563 GCGCGGTGGAGCAACAGTCTCAACATCGCCCAAGGAGTGCATCAGGTGGCGGCC 7622
 QY 198 IleLeuAsnHis 201
 Db 7623 GTGCTCAACAC 7634
 RESULT 10
 ABQ82346
 ID ABQ82346 standard; cDNA; 8487 BP.
 AC ABQ82346;
 XX
 XX 17-DEC-2002 (first entry)
 DE Human NOV15d encoding cDNA SEQ ID NO:41.
 XX
 KW Human; NOVX: cytostatic; neuroprotective; anticonvulsant; cardiovascular;
 KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
 KW antihelminthic; antirheumatic; immunosuppressive; antiallergic; virucide;
 KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
 KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
 KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
 KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
 KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
 KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
 KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
 KW rheumatoid arthritis; gene; chromosome 4; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 299..8140
 FT /*tag= a
 FT /product= "NOV15d"
 XX
 XX WO200262999-A2.
 XX
 XX 15-AUG-2002.
 XX
 XX 31-DEC-2001; 2001WO-0049976.
 XX
 XX 29-DEC-2000; 2000US-0258928P.
 PR 02-JAN-2001; 2001US-0259415P.
 PR 04-JAN-2001; 2001US-0259785P.
 PR 20-FEB-2001; 2001US-0269814P.
 PR 09-MAR-2001; 2001US-0279863P.
 PR 29-MAR-2001; 2001US-0279832P.
 PR 29-MAR-2001; 2001US-0279833P.
 PR 13-APR-2001; 2001US-0283899P.
 PR 18-APR-2001; 2001US-028447P.
 PR 25-APR-2001; 2001US-0286683P.
 PR 25-MAY-2001; 2001US-0294080P.
 PR 16-AUG-2001; 2001US-0312915P.
 PR 17-AUG-2001; 2001US-0313325P.
 PR 17-SEP-2001; 2001US-0322699P.
 PR 26-NOV-2001; 2001US-0333350P.

XX PA (CURA-) CURAGEN CORP.

XX PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;

XX PI Malyankar U, Shinkets RA, Tchernev VT, Spaderna SK, Gorman L;

XX PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;

XX PI Rastelli L, Casman SJ, Boldog F, Burgess CB, Edinger S, Ellerman K;

XX PI Gunther E, Smithson G, Millet I, Macdougall JR;

XX DR WPI: 2002-732706/79.

XX DR P-PSDB; ABP3589.

XX DR New NOVX polypeptides and polynucleotides useful for treating NOVX-

XX DR associated disorders, such as cancers, neurological disorders, disorders

XX DR of vesicular transport, gastrointestinal disorders, and autoimmune

XX DR diseases.

XX PS Claim 8; Page 123-125; 444pp; English.

XX CC The present invention describes novel human proteins designated NOVX,

XX CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,

XX CC cytotatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,

XX CC antidiabetic, antiinflammatory, antirheumatic, antiallergic, virucide,

XX CC immunosuppressive, antiallergic, antianemic, antibacterial, fungicide,

XX CC protozoacide and antihelminthic activities, and can be used in gene

XX CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the

XX CC manufacture of a medicament for treating a syndrome associated with a

XX CC human disease selected from NOVX-associated disorder, such as cancers

XX CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,

XX CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,

XX CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's

XX CC disease), disorders of vesicular transport (e.g. cystic fibrosis,

XX CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal

XX CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),

XX CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic

XX CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic

XX CC and protozoal infections. The NOVX proteins can be used as immunogens to

XX CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be

XX CC used in chromosome mapping, identifying individuals from minute

XX CC biological samples (tissue typing), and in forensic identification of a

XX CC biological sample. The present sequence encodes human NOV15d, which is

XX CC located on chromosome 4

XX SQ Sequence 8487 BP; 2306 A; 2029 C; 2148 G; 2001 T; 0 U; 3 Other;

Alignment Scores:

| | | | |
|------------------------|---------|---------------|------|
| Pred. No.: | 1,4e-58 | Length: | 8487 |
| Score: | 566.00 | Matches: | 107 |
| Percent Similarity: | 71.08% | Conservative: | 38 |
| Best Local Similarity: | 52.45% | Mismatches: | 47 |
| Query Match: | 52.55% | Indels: | 12 |
| DB: | 6 | Gaps: | 4 |

US-10-029-020-14_COPY_2400_2600 (1-201) x ABQ82346 (1-8487)

QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20

Db 7046 ATTGGATTTCATGGTGGCCCTGTATGACCCACTCACCAATTAATCCATTGGAGAA 7105

QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGlyLeuTyrLysHisLeu 40

Db 7106 GATTATGACATTTCGGCAGACGGTGCACACACCTGCATAGAAATCTGGAAAGAAT 7165

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60

Db 7166 ---GGGAAGGACCCAGCTCTTTTAACTGTACATGTTTGAATAACAACCTCGCAAGC 7222

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80

Db 7223 AAAATCCATGACGGTGAAGAGATTATCATCAGATGTTTACAGCTGGCTGGTGGATTTGGT 7282

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100

Db 7283 TTCATCTGCACAAATGCTATTCTCGATTCCCTGTTCCCAATTTTGAATTTACAGAACCT 7342

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120

Db 7343 TCTTACGAACCTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 7390

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140

Db 7391 TTCGAGAGTCCAGCAGCAAGTGGCGCGCAGCCAGGCGCTTCTCTGCTGGGGAAGATG 7450

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159

Db 7451 GCCAGGTGCAG-----GTGAGCGCGCGCGCGCGCGCGCGCTCTCTGG 7498

QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177

Db 7499 CTGTGGTTCGCCAGGTCAAGTCGTCATCGCAAGGCGCGTCATCGTGGCGCTCAGCCAG 7558

QY 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAla 197

Db 7559 GCGCGCGTGCAGACCAACGCTCAACATCGCCCAACGAGGACTGCATCAAGTGGCGGCC 7618

QY 198 IleLeuAsnHis 201

Db 7619 GTGCTCAACAC 7630

RESULT 11

ABQ82344

ID ABQ82344 standard; cDNA; 8645 BP.

XX AC ABQ82344;

XX DT 17-DEC-2002 (first entry)

XX DE Human NOV15b encoding cDNA SEQ ID NO:37.

XX KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;

XX KW cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;

XX KW antirheumatic; antiallergic; immunosuppressive; antiallergic; virucide;

XX KW antianemic; antibacterial; protozoacide; antihelminthic; gene therapy;

XX KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;

XX KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; goitre;

XX KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;

XX KW diabetes mellitus; Grave's disease; gastrointestinal disorder; infection;

XX KW ulcerative colitis; gastric disorder; duodenal disorder; autoimmune

XX KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;

XX KW rheumatoid arthritis; gene; chromosome 4; ss.

XX OS Homo sapiens.

XX FH Key

XX CDS 151..8316

XX FT /*tag= a

XX FT /product= "NOV15b"

XX PN W0200262999-A2.

XX PD 15-AUG-2002.

XX XX 31-DEC-2001; 2001WO-US049976.

XX XX 29-DEC-2000; 2000US-0258928P.

XX XX 02-JAN-2001; 2001US-0259415P.

XX XX 20-FEB-2001; 2001US-0269814P.

XX XX 09-MAR-2001; 2001US-0279863P.

XX XX 29-MAR-2001; 2001US-0279832P.

XX XX 13-APR-2001; 2001US-0283889P.

XX XX 18-APR-2001; 2001US-0284447P.

XX XX 25-MAY-2001; 2001US-0286683P.

XX XX 29-MAY-2001; 2001US-0294080P.

XX XX 16-AUG-2001; 2001US-0312915P.

PR 17-AUG-2001; 2001US-0313325P.
 PR 17-SEP-2001; 2001US-0322699P.
 PR 26-NOV-2001; 2001US-0333350P.
 XX (CURA-) CURAGEN CORP.
 XX
 XX
 PI Spytek KA, Li L, Wolenc AR, Vernet CM, Eisen A, Liu X;
 PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
 PI Kekuda R, Patturajan M, Gusev F, Gangolli EA, Guo X, Shenoy S;
 PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
 PI Gunther E, Smithson G, Millet I, Macdougall JR;
 XX WPI; 2002-732706/79.
 DR P-PSDB; ABP53587.
 XX
 XX
 PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
 PT associated disorders, such as cancers, neurological disorders, disorders
 PT of vesicular transport, gastrointestinal disorders, and autoimmune
 PT diseases.
 XX
 PS Claim 8; Page 114-117; 444pp; English.
 XX
 CC The present invention describes novel human proteins designated NOVX,
 CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
 CC cytotstatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,
 CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
 CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
 CC protozoacide and antihelminthic activities, and can be used in gene
 CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease selected from NOVX-associated disorder, such as cancers
 CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
 CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
 CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
 CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
 CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
 CC disorders (e.g. ulcerative colitis, or gastritis, and duodenal disorders),
 CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
 CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
 CC and protozoal infections. The NOVX proteins can be used as immunogens to
 CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
 CC used in chromosome mapping, identifying individuals from minute
 CC biological samples (tissue typing), and in forensic identification of a
 CC biological sample. The present sequence encodes human NOV15b, which is
 CC located on chromosome 4
 XX
 SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;
 Alignment Scores:
 Pred. No.: Length: 8645
 Score: 1.43e-58 Matches: 107
 Percent Similarity: 71.08% Conservative: 38
 Best Local Similarity: 52.45% Mismatches: 47
 Query Match: 52.55% Indels: 12
 Ds: 6 Gaps: 4
 US-10-029-020-14_COPY_2400_2600 (1-201) x ABQ82344 (1-8645)
 QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
 Db 7222 ATTGATTTCATGGTGGCTGATGACCCATCCACCAATAATTAATCCACTTTGGAGAAAGA 7281
 QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrPlyHisLeu 40
 Db 7282 GATTATGACATTTTGGAGGAGCGTGGACACACCTGACATAGAAATCTGGAAGAAGAAATT 7341
 QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProPheSer 60
 Db 7342 ---GGGAAGGACCCAGCTCCCTTTAACTTGTACATGTTTAGAATAACAAACCTGCAAGC 7398
 QY 61 AsnSerGlnAspIleLeuGlyCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
 Db 7399 AAATCCATGACGCGAAAGATTATCATCATCAGATGTTTACAGCTGGCTGGTGCATTTGGT 7458
 QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 Db 7459 TTCCATCTGCACATGCTATTCTCTGATTCCTTCCAAATTTGATTAAACAGAACCT 7518
 QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
 Db 7519 TCTTACGAACCTTGTG-----AAGAGTCAGCAGTGGATGATATACCGCCCATC 7566
 QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 Db 7567 TTCGAGTCCAGCAGCAAGTCGCGCAGCCAGGCTTCCTGCTCGCGGAGAGATG 7626
 QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
 Db 7627 GCCGAGGTGCAG-----GTAGCCGCGCGCGCGCGCGCGCGCGCGCTCTCGG 7674
 QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
 Db 7675 CTGTGGTTCGCCACGTCATCGCAAGGCGCTCATCTGCGCTGCGCTGAGCCAG 7734
 QY 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAla 197
 Db 7735 GCGCGGTGCGAGCAACGTCCTCAACATCGCCCAACGAGGACTGCATCAAGGTGCGCGCC 7794
 QY 198 IleLeuAsnHis 201
 Db 7795 GTGCTCAACAAC 7806
 RESULT 12
 ABQ82343
 ID ABQ82343 standard; cDNA; 8675 BP.
 AC ABQ82343;
 XX
 DT 17-DEC-2002 (first entry)
 DE Human NOV15a encoding cDNA SEQ ID NO:35.
 KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
 KW cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;
 KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
 KW antianaemic; antibacterial; protozoacide; antihelminthic; Gene therapy;
 KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
 KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; Pick's
 KW disease; vesicular transport disease; cystic fibrosis; goitre;
 KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
 KW ulcerative colitis; gastritis; duodenal disorder; infection;
 KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
 KW rheumatoid arthritis; gene; chromosome 4; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 CDS 151..8328
 FT /*tag= a
 FT /product= "NOV15a"
 FT /transl_except= (pos.1249..1251,aa:Ser)
 WO200262999-A2.
 15-AUG-2002.
 31-DEC-2001; 2001WO-US049976.
 29-DEC-2000; 2000US-0258928P.
 02-JAN-2001; 2001US-0259415P.
 04-JAN-2001; 2001US-0259785P.
 20-FEB-2001; 2001US-0269814P.
 09-MAR-2001; 2001US-0279863P.
 29-MAR-2001; 2001US-0279832P.
 29-MAR-2001; 2001US-0279833P.
 13-APR-2001; 2001US-0283889P.

18-APR-2001; 2001US-0284447P.
 25-APR-2001; 2001US-0286683P.
 29-MAY-2001; 2001US-0294080P.
 16-AUG-2001; 2001US-0312915P.
 17-AUG-2001; 2001US-0313225P.
 17-SEP-2001; 2001US-0322699P.
 26-NOV-2001; 2001US-0333350P.
 (CURA-) CURAGEN CORP.
 Spytok KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
 PI Malyankar U, Shimkets RA, Tchernev VT, Spaderina SK, Gorman L;
 PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
 PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
 PI Gunther E, Smithson G, Millet I, Macdougall JR;
 XX WPI; 2002-732706/79.
 DR P-PSDB; ABP53586.
 XX
 XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
 associated disorders, such as cancers, neurological disorders, disorders
 of vesicular transport, gastrointestinal disorders, and autoimmune
 diseases.
 XX
 XX Claim 8; Page 110-112; 444pp; English.
 XX
 XX The present invention describes novel human proteins designated NOVX,
 where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
 cytoskeletal, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
 antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
 immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
 protozoacide and antihelminthic activities, and can be used in gene
 therapy. The NOVX proteins, nucleotides or antibodies can be used in the
 manufacture of a medicament for treating a syndrome associated with a
 human disease selected from NOVX-associated disorder, such as cancers
 (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
 ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
 ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
 disease), disorders of vesicular transport (e.g. cystic fibrosis,
 diabetes mellitus, Grave's disease, or goitre), gastrointestinal
 disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
 autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
 anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
 and protozoal infections. The NOVX proteins can be used as immunogens to
 produce antibodies and as vaccines. The NOVX nucleotide sequences may be
 used in chromosome mapping, identifying individuals from minute
 biological samples (tissue typing), and in forensic identification of a
 biological sample. The present sequence encodes human NOV15a, which is
 located on chromosome 4
 XX
 XX Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;
 SQ
 Alignment Scores:
 Pred. No.: 1-44e-58 Length: 8675
 Score: 566.00 Matches: 107
 Percent Similarity: 71.08% Conservative: 38
 Best Local Similarity: 52.45% Mismatches: 47
 Query Match: 12 Indels: 4
 DB: 6 Gaps: 4
 US-10-029-020-14_COPY_2400_2600 (1-201) x ABQ82343 (1-8675)
 QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrIleValHisMetGlyArgArg 20
 DB 7234 ATTGGATTTCATGGTGGCTGTATGACCCACTCACCACAAATTAATCCACTTGGAGAAAGA 7293
 QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
 DB 7294 GATTATGACATTTTGGCAGACGGTGGACACCTGACATAGACATAGAAATCTGGAAAGAATT 7353
 QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
 DB 7354 ---GGGAGGAGCCAGCTCCTTTTAACTGTGTACATGTTTGGATAACAACCTGCAAGC 7410

61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
 7411 AAAATCCATGACGTGAAGATTATACATCAGATGTTAAACAGCTGGGTGGTACATTGGT 7470
 QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
 DB 7471 TTCCATCTGCACAATGCTATTCTCGATTCCCTGTTCCCAAAATTTGATTTAACAGAACCT 7530
 QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
 DB 7531 TCTTACGACTGTG-----AAGATCAGCAGTGGGATGATATACCCGCCCATC 7578
 QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
 DB 7579 TTCGAGTCCAGCAGCAAGTGGCGCGCAGCCCAAGGCTTCTCTGCTGGGAAAGATG 7638
 QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
 DB 7639 GCCAGGTGCAG-----GTGAGCCGCGCGCGCGCGCGCGCGCGCTCTGG 7686
 QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
 DB 7687 CTGTGTTTCCGACCGTCAAGTCTGATCGCAAGGCGCTCATGCTGCGCTCAGCCAG 7746
 QY 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAla 197
 DB 7747 GGCGCGTGCAGACCAACGCTCTCAACATCGCCACGAGGACTGCATCAAGGTGCGCGCC 7806
 QY 198 IleLeuAsnHis 201
 DB 7807 GTGCTCAACAAC 7818
 RESULT 13
 ABK34316
 ID ABK34316 standard; cDNA; 791 BP.
 XX
 XX ABK34316;
 AC
 XX 08-MAY-2002 (first entry)
 DT
 DE Human cDNA for novel secreted protein, SEQ ID 85.
 XX
 XX Human; ss; gene; secreted protein; immune deficiency; viral infection;
 KW bacterial infection; fungal infection; autoimmune disorder; burn;
 KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
 KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
 KW coagulation disorder; haemophilia; inflammatory disorder; ulcer;
 KW tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
 KW lymphoid cell deficiency.
 XX
 XX Homo sapiens.
 OS
 XX WO200177290-A2.
 XX
 XX 18-OCT-2001.
 XX
 XX 29-MAR-2001; 2001WO-US010295.
 XX
 XX 06-APR-2000; 2000US-0194941P.
 XX
 XX (GEMY) GENETICS INST INC.
 PA
 XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Gulkota K, Graham JR;
 XX
 XX WPI; 2002-179323/23.
 XX
 XX Six hundred and twenty five polynucleotides derived from a variety of
 human tissue sources which encode such as autoimmunes, useful for treating
 immune deficiencies and disorders such as autoimmune disorders.
 XX

P-PSDB; ABR59318.

New isolated human differentially-regulated breast cancer polynucleotide
and polypeptide, useful for diagnosing, staging, prognosticating,
preventing and/or treating diseases and conditions relating to breast
cancer.

XX Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.
XX

The invention relates to isolated polynucleotides which are
differentially-regulated in breast cancer. The methods and compositions
of the present invention are useful for detecting, diagnosing, staging,
monitoring, prognosticating, preventing and/or treating diseases and
conditions relating to breast cancer, and may be used in gene therapy or
antisense therapy. They can also be used in research, drug discovery,
clinical medicine and forensic medicine. Sequences given in records,
ACC72012-ACC72074 represent polynucleotides of the invention that are
differentially-regulated in breast cancer. NOTE: The sequence data for
this patent did not form part of the printed specification, but was
obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

XX Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 5.63e-54 Length: 9695
Score: 529.50 Matches: 103
Percent Similarity: 68.78% Conservative: 38
Best Local Similarity: 50.24% Mismatches: 49
Query Match: 49.16% Indels: 15
DB: Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x ACC72052 (1-9695)

QY 1 IleGlyTyrHisGlyCysLeuThrAspProLeuThrLysLeuValHisMetGlyArg 20
Db 7236 ATTGGCTTCATGGGGAGCTCATGACCCCTTGACCAAGCTGGTCCACTTCACTCAGCGT 7295
QY 21 AspTyrAspValLeuAlaGlyArgTrpSerProAspHisGluLeuThrPylHisLeu 40
Db 7296 GATTATGATGTGTCGCAGGAGCATGGACTCCCCAGACTATACCATGTGGAAAACGTG 7355
QY 41 SerSerSerAsnValMetProPheAsnLeuTyMetPheLysAsnAsnAsnProlleSer 60
Db 7356 GGCAAGGAGCGGCC---CCCCTTAACCTGTATATGTTCAAAGACAAACAATCTCTCAGC 7412
QY 61 AsnSerGlnAspileLysCysPheMetThrAspValAsnSertPrleLeuThrPheGly 80
Db 7413 AGTAGCTAGATTGAAGAACIACGTGCAGATGTGAAAAGCTGGCTGTGATGTTTTGA 7472
QY 81 PheGlnLeuHisasnVallileProGlyTyrrProLysProaspMetAspalametGlupro 100
Db 7473 TTTCAGCTTAGCAACATCATTCCTGGCTGCCGAGAGCAATAATGATTTTCGGCTCCCT 7532
QY 101 SeryrGluLeulleHisThrGlnMetLysThrglnLuTrpAspnSerlysSerille 120
Db 7533 CCCTATGAATTGTCAGAGAGTCAAGCAAGT-----CAGATGCAGACGCTCATT 7580
QY 121 LeuglyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7581 ACAGGTGTCCAACAGACAGACAGACAGACATAAACCAGGCCCTTCATGGCTCFGAA----- 7634
QY 141 AspGlnLeutyryglyserThrilleTherSer-CysglndlnAlaprollysThrlylsys-- 159
Db 7635 -----GGACAGGTCATTACTAAAGAGCTCCACGCCAGCATCCGAGAGAAAAAGCA 7682
QY 160 -----PheAlaserSerglyserValpheelylysglyvallyspheAlaleuLys 176
Db 7682 GGTCACCTGGITTCACCAACCACGCCCATCATTTGGCAAAGGCATCATGTTTGGCATCAA 7742
QY 177 AspGlyArgValThrAspIlelleSerValalaenGluAspGlyArgValala 196
Db 7743 GAAGGGCGGGTGACACACGGGGCGGTTCACAGATCCCGACGAGATAGCCGCAAGTGGCA 7802

OY 197 AlaIleLeuAsnHis 201
Db :::::|||||:::
7803 TCTGTGCTGACAAAC 7817

Search completed: August 14, 2004, 02:34:36
Job time : 386.95 secs

This Page Blank (uspto)

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 1084.27 Seconds
 (without alignments)
 2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_1760_2300
 Perfect score: 2845
 Sequence: 1 YIIGADGSLRLLLANGMEVA.....YNRAGSWVRVYDGLGRV 541

Scoring table:

| | |
|---------------------------|--|
| BLOSUM62 | |
| Xgapop 10.0 , Xgapext 0.5 | |
| Ygapop 10.0 , Ygapext 0.5 | |
| Fgapop 6.0 , Fgapext 7.0 | |
| Delop 6.0 , Delext 7.0 | |

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Command line parameters:
 -MODEL=frame+ p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO.spool/US1002020/runat_06082004_112218_29331/app_query.fasta_1.3519
 -DB=Published Applications NA_QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
 -LOOPCL=0 -LOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
 -MAXLEN=2000000000 -USER=US1002020@cgn1.1.2156 grunat_06082004_112218_29331
 -NCFU=6 -ICPU=3 -NO MAP -LARGQUEY -NEG_SCORES=0 -WAIT -DSPELOCK=100
 -LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

| | |
|-----|---|
| 1: | /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: |
| 2: | /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq: |
| 3: | /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq: |
| 4: | /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq: |
| 5: | /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: |
| 6: | /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq: |
| 7: | /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: |
| 8: | /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq: |
| 9: | /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq: |
| 10: | /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq: |
| 11: | /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq: |
| 12: | /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq: |
| 13: | /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2: |
| 14: | /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq: |
| 15: | /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq: |
| 16: | /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq: |
| 17: | /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq: |
| 18: | /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq: |
| 19: | /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: |

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
|------------|-------|-------------|--------|----|-------------|

ALIGNMENTS

RESULT 1

US-10-383-201-43
 ; Sequence 43, Application US10383201
 ; Publication No. US20040029226A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alsbrook II, John et al.
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
 ; FILE REFERENCE: 21402-568A
 ; CURRENT APPLICATION NUMBER: US/10/383,201
 ; PRIOR FILING DATE: 2003-03-06
 ; PRIOR APPLICATION NUMBER: 10/029020
 ; PRIOR FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/365,984
 ; PRIOR FILING DATE: 2002-03-20
 ; PRIOR APPLICATION NUMBER: 60/372,022
 ; PRIOR FILING DATE: 2002-04-12
 ; PRIOR APPLICATION NUMBER: 60/389,143
 ; PRIOR FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 60/391,779
 ; PRIOR FILING DATE: 2002-06-26
 ; PRIOR APPLICATION NUMBER: 60/410,755
 ; PRIOR FILING DATE: 2002-09-13
 ; PRIOR APPLICATION NUMBER: 60/412,957

| | | | | | | |
|----|--------|-------|-------|----|----------------------|-------------------|
| 1 | 2845 | 100.0 | 8354 | 13 | US-10-383-201-43 | Sequence 43, Appl |
| 2 | 2845 | 100.0 | 8354 | 13 | US-10-029-020-13 | Sequence 13, Appl |
| 3 | 2821.5 | 99.2 | 8438 | 13 | US-10-042-865-1 | Sequence 1, Appl |
| 4 | 2751 | 96.7 | 8355 | 13 | US-10-383-201-55 | Sequence 55, Appl |
| 5 | 2233 | 78.5 | 3614 | 13 | US-10-342-887-1743 | Sequence 1743, Ap |
| 6 | 2233 | 78.5 | 3614 | 13 | US-10-172-118-1743 | Sequence 39, Appl |
| 7 | 2233 | 78.5 | 8473 | 17 | US-10-038-854-39 | Sequence 39, Appl |
| 8 | 2233 | 78.5 | 8473 | 17 | US-10-038-854-41 | Sequence 37, Appl |
| 9 | 2233 | 78.5 | 8645 | 17 | US-10-038-854-37 | Sequence 35, Appl |
| 10 | 2233 | 78.5 | 8675 | 17 | US-10-038-854-35 | Sequence 74, Appl |
| 11 | 2141.5 | 75.3 | 8797 | 9 | US-09-808-602-74 | Sequence 77, Appl |
| 12 | 2141.5 | 75.3 | 8797 | 9 | US-09-808-602-77 | Sequence 62, Appl |
| 13 | 2141.5 | 75.3 | 8797 | 10 | US-09-800-198-62 | Sequence 65, Appl |
| 14 | 2141.5 | 75.3 | 8797 | 10 | US-09-800-198-65 | Sequence 79, Appl |
| 15 | 2138.5 | 75.2 | 8409 | 9 | US-09-808-602-79 | Sequence 76, Appl |
| 16 | 2138.5 | 75.2 | 8409 | 9 | US-09-800-198-67 | Sequence 67, Appl |
| 17 | 2136.5 | 75.1 | 6560 | 10 | US-09-808-602-76 | Sequence 64, Appl |
| 18 | 2136.5 | 75.1 | 6560 | 10 | US-09-800-198-64 | Sequence 79, Appl |
| 19 | 2136.5 | 75.1 | 9058 | 16 | US-10-144-194A-79 | Sequence 81, Appl |
| 20 | 2136.5 | 75.1 | 9695 | 16 | US-10-144-194A-81 | Sequence 12, Appl |
| 21 | 2136.5 | 75.1 | 9729 | 9 | US-09-808-602-12 | Sequence 12, Appl |
| 22 | 2136.5 | 75.1 | 9729 | 10 | US-09-800-198-12 | Sequence 7, Appl |
| 23 | 2136.5 | 75.1 | 9826 | 9 | US-09-808-602-7 | Sequence 7, Appl |
| 24 | 2136.5 | 75.1 | 9826 | 10 | US-09-800-198-7 | Sequence 78, Appl |
| 25 | 2135.5 | 75.1 | 8689 | 9 | US-09-808-602-78 | Sequence 66, Appl |
| 26 | 2135.5 | 75.1 | 8689 | 10 | US-09-800-198-66 | Sequence 143, App |
| 27 | 2106.5 | 74.0 | 8575 | 13 | US-10-072-012-143 | Sequence 927, App |
| 28 | 2048.5 | 72.0 | 12880 | 16 | US-10-295-027-927 | Sequence 21273, A |
| 29 | 982.5 | 34.5 | 806 | 9 | US-09-864-761-21273 | Sequence 4526, Ap |
| 30 | 982.5 | 34.5 | 1207 | 15 | US-10-198-846-9717 | Sequence 13976, A |
| 31 | 982.5 | 34.5 | 1973 | 9 | US-09-864-761-4526 | Sequence 6607, Ap |
| 32 | 735.5 | 25.9 | 648 | 15 | US-10-198-846-7548 | Sequence 104, App |
| 33 | 665 | 23.4 | 3217 | 15 | US-10-198-846-13976 | Sequence 15932, A |
| 34 | 520 | 18.3 | 330 | 11 | US-09-864-408A-6607 | Sequence 34930, A |
| 35 | 387 | 13.6 | 286 | 14 | US-10-040-739-104 | Sequence 3302, Ap |
| 36 | 369 | 13.0 | 307 | 15 | US-10-029-386-15932 | Sequence 11416, A |
| 37 | 369 | 13.0 | 512 | 15 | US-10-029-386-2232 | Sequence 17002, A |
| 38 | 248.5 | 8.7 | 3957 | 16 | US-10-369-493-34930 | Sequence 32730, A |
| 39 | 247 | 8.7 | 586 | 15 | US-10-029-386-3302 | Sequence 32058, A |
| 40 | 240.5 | 8.5 | 4545 | 13 | US-10-282-122A-11416 | Sequence 3837, Ap |
| 41 | 240 | 8.4 | 245 | 15 | US-10-029-386-17002 | Sequence 33231, A |
| 42 | 219 | 7.7 | 3759 | 13 | US-10-282-122A-32730 | |
| 43 | 212 | 7.5 | 4155 | 13 | US-10-282-122A-32058 | |
| 44 | 205 | 7.2 | 5394 | 16 | US-10-398-221-3837 | |
| 45 | 204.5 | 7.2 | 4716 | 13 | US-10-282-122A-33231 | |

Mon Aug 16 09:01:11 2004

;; PRIOR FILING DATE: 2002-09-23
;; PRIOR APPLICATION NUMBER: 10/051,874
;; PRIOR FILING DATE: 2002-01-16
;; PRIOR APPLICATION NUMBER: 60/366,928
;; PRIOR FILING DATE: 2002-03-22
;; PRIOR APPLICATION NUMBER: 10/055,877
;; PRIOR FILING DATE: 2002-01-22
;; NUMBER OF SEQ ID NOS: 155
;; SOFTWARE: Curaseq1ist version 0.1
;; SEQ ID NO 43
;; LENGTH: 8354
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 0 Length: 8354
Score: 2845.00 Matches: 541
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-383-201-43 (1-8354)

QY 1 TyrTyrIleGlyValaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
DB 5312 TACTACATCGGGCCGATGGCTCTTGGCGCTGCTGGCCAAACGGCATGGAGTGGCG 5371
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
DB 5372 CTGCAGACTGAGCCCACTTGTGGTGGCCACCGTCAACCCCAACCGTGGCAAGGAAT 5431
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
DB 5432 GTCAACGCTGCCATCGACACGGCCCTCACTCTTGGCGCGCGCTGGCGTGCACAAACCGAAATCTCCTA 5551
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
DB 5492 GCTCGGGCCGCGTCACTGTCTTGGCGCGCGCTGGCGTGCACAAACCGAAATCTCCTA 5551
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
DB 5552 TCTCTGGACTTGTATCGGTAAACACGACACGACGAGAAAGATCTATGATCACCCCGCAAGTTC 5611
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
DB 5612 ACCCTTCGGGATTCGTACGACCGAGCGGGCGGGCGCCAGCTCTGGTCAACCCAGCAGCAGG 5671
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
DB 5672 CTGATGGTGTCAACGTGACATCTCCCTGGGGGTTACATTTGGTCAATCCAGAGGGC 5731
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
DB 5732 ATCATGCTGAAAGATGAATACGACGAGCGGGCGGCGATCATCATCAGGATCTTCGCT 5791
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
DB 5792 GATGGGAAGACATGGAGCTACACATCTTAGAAGTCCATGGTGTCTACTACACAGC 5851
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
DB 5852 CAGAGGCAGTATATCTTTGAGTTCGACAAAGATACCGCCCTCTCTCTGTGACGATGCC 5911
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
DB 5912 AACGTGGCGCGACACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTAT 5971
QY 221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

DB 5972 CAGCCCCCTGAGGCAATGCTCAGTCATACAGGACTTCTACAGGATGGGCACCTCCTT 6031
QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
DB 6032 CACACCTTCTACCTGGGCACCTGGCGGAGGTGATATACAAGTATGGCAAACTGTCAAAG 6091
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrLagly 280
DB 6092 CTGGCAGAGACGCTCTATGACACCAAGGTCACTTTCACCTATGACGAGAGCGCAGGC 6151
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
DB 6152 ATGCTGAAGACCATCAACCTACAGAAATGAGGCTTCCCTGACCATTCGCTACCGTCTAG 6211
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
DB 6212 ATTGGGCCCCGTGATTGACCGACAGATCTTCGCTTCACTGAGGAAGGATGGTCAACGCC 6271
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
DB 6272 CGTTTTGACTACAACTATGACAAACAGCTTCCGGGTGACAGCATGCGGCTGTGATCAAC 6331
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
DB 6332 GAGACCCCACTGCCATTTGATGCTATGATGATGTTGTCAGGCAAGACAGACGACG 6391
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
DB 6392 TTTGGGAAGTTTGGTGTCTATTACTATGACATTAACAGATCATCACACAGCTGTCTATG 6451
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
DB 6452 ACCCACACCAAGCATTTGATGCTATGCGAGATGAGGAAGTGCAGTATGATGATCTTC 6511
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
DB 6512 CGCTCGCTCATGTACTGGGATGACCCGTCAGTATGATTAACATGCGGCGAGTAGTGAAGAG 6571
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
DB 6572 GAGCTGAGGTAGGACCTTACGCCAATACCACTCGCTACTCTCTATGATGATGATGCTAC 6631
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
DB 6632 GGCAGCTGCAGACAGTCTCCATCAATGACAAACCCACTCTGGCGCTACAGCTAGACCTC 6691
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
DB 6692 AATGGGAACCTGCATTTACTGAGCCCTGGGACAGTGCAGCGCTCACACCATACGCTAT 6751
QY 481 AspileArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
DB 6752 GACATCGCGCCGCGCATCACTCGCTGGGTGACGTGCAATACAAGATGGATGGAGTGGC 6811
QY 501 PheLeuArgGlnArgGlyGlyAspilePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
DB 6812 TTCTGAGGCGAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCTGCTCATCAAG 6871
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
DB 6872 GCCTTACAAACCGGGCTGGCAGCTGGAGTGCAGGTACCGCTACGATGGCTGGCGCGCGC 6931
QY 541 Val 541
DB 6932 GTG 6934

RESULT 2
US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.

! TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same

! FILE REFERENCE: 21402-225
 ! CURRENT APPLICATION NUMBER: US/10/029,020
 ! CURRENT FILING DATE: 2001-12-19
 ! PRIOR APPLICATION NUMBER: 60/256,704
 ! PRIOR FILING DATE: 2000-12-19
 ! PRIOR APPLICATION NUMBER: 60/311,590
 ! PRIOR FILING DATE: 2001-08-10
 ! PRIOR APPLICATION NUMBER: 60/257,314
 ! PRIOR FILING DATE: 2000-12-20
 ! PRIOR APPLICATION NUMBER: 60/311,613
 ! PRIOR FILING DATE: 2001-08-10
 ! PRIOR APPLICATION NUMBER: 60/315,617
 ! PRIOR FILING DATE: 2001-08-29
 ! PRIOR APPLICATION NUMBER: 60/307,506
 ! PRIOR FILING DATE: 2001-07-24
 ! PRIOR APPLICATION NUMBER: 60/322,358
 ! PRIOR FILING DATE: 2001-09-14
 ! PRIOR APPLICATION NUMBER: 60/294,075
 ! PRIOR FILING DATE: 2001-05-29
 ! PRIOR APPLICATION NUMBER: 60/288,153
 ! PRIOR FILING DATE: 2001-05-02
 ! NUMBER OF SEQ ID NOS: 190
 ! SOFTWARE: PatentIn Ver. 2.1
 ! SEQ ID NO 13
 ! LENGTH: 8354
 ! TYPE: DNA
 ! ORGANISM: Homo sapiens
 ! US-10-029-020-13

Alignment Scores:

Pred. No.: 0 Length: 8354
 Score: 2845.00 Matches: 541
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-029-020-13 (1-8354)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
 DB 5312 TACTACATCGGGCGCGATGGCTCCTTCGGCTGCTGCTGGCCAAACCGCGAGGTGGCG 5371
 QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
 DB 5372 CTCGAGACTGAGCCCACTTCTGGCTGGCGACCGCTCAACCCCGGGGCAAGAGAAAT 5431
 QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
 DB 5432 GTCACGCTGCCCATCGACAAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGCAG 5491
 QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
 DB 5492 GCTCGGGGCGAGGTCACTGTCTTTGGGCGCGCTGCGGGTGACCAACCGAAATCTCCTA 5551
 QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
 DB 5552 TCTCTGAGCTTTGATCGCGTAACCGCAACGACGAGAGAGATCATGATGACCAACCGCAAGTTC 5611
 QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
 DB 5612 ACCCTTCGGATTCTGTACGACACGCGGGCGCGCCAGCTCTGGTCAACCCAGCAGCAGG 5671
 QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
 DB 5672 CTGAATGGTGTCAACGTGACATACCTCCCTGGGGGTTCATTGCTGGCATCCAGAGGGCG 5731
 QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
 DB 5732 ATCATGCTCTGAAGATGGATACGACACGAGCGGGCGCGCATCATCATCCAGGATCTTCGCT 5791
 QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180

DB 5792 GATGGAGACATGAGGCTACACATCTTAGAAGTCCATGGTCTGTCTACTACAGC 5851
 QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
 DB 5852 CAGAGCGAGTATATCTTTGAGTTCCAGAAGATACCGGCTCTCTTCTGTGAGATGCC 5911
 QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
 DB 5912 AACGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTGGGTACTACAGAACAATCTAT 5971
 QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
 DB 5972 CAGCCCCCTGAGGGCAATGCCCTCAGTCATACAGGACTTCACTAGGATGGGCACTCCT 6031
 QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
 DB 6032 CACACCTTCTACCTGGGCACTGGCGCAGGGTGATATCAAGTATGGCAAACTGTCAAAG 6091
 QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
 DB 6092 CTGGCAGAGACGCTCTATGACACCAACCAAGGTCACTTCACTATGACAGACGCGCAGC 6151
 QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
 DB 6152 ATGCTGAAGACCATCAACCTACAGATGAGAGGCTTCACTGCACCATCCGCTACCGTCAG 6211
 QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
 DB 6212 ATTGGCCCCGTGATTGACCGACAGATCTTCGCTTCACTGAGGAAGGATGGTCAACGCC 6271
 QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
 DB 6272 CGTTTTCAGCTACAACTATGACCAACAGCTTCGGGTGACCAAGTGCAGGCTGTGATCAAC 6331
 QY 341 GlnThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
 DB 6332 GAGACCCCATGGCCCATTTGATCTCTATCGCTATGATGTGTGAGCAAGACAGAGCAG 6391
 QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleThrThrAlaValMet 380
 DB 6392 TTTGGGAAGTTTGGTGTCACTTACTATGACATTAACAGATCATCATCAGCTGTGTCATG 6451
 QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
 DB 6452 ACCCACCAAGCATTTTGATGCATATGGCAGCATGAAGCAAGTGCAGTATGAGATCTTC 6511
 QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
 DB 6512 CGTCTGCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGGGAGTAGTGAAGAAG 6571
 QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
 DB 6572 GAGCTGAAGGTAGGACCTTACGCCAATACCACTCGCTACTCTCTATGAGTATGATGCTGAC 6631
 QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
 DB 6632 GGCCAGCTCGACAGACTCTCCATCAATGACCAAGCCACTCTGGCGCTACAGTACGACCTC 6691
 QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
 DB 6692 AATGGGAACCTCGACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCATACGGTAT 6751
 QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
 DB 6752 GACATCCGGACCGCATCTCTGGCTGGGTGACGTGCAATACAAAGATGATGAGGATGGC 6811
 QY 501 PheLeuArgGlnArgGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
 DB 6812 TTCCTGAGCAGCGGGCGGTGATATCTTTGAGTCAACCTACGCTGGCTGTCTCATCAAG 6871
 QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540

QY 278 ThrAlaGlyMetLeuLeuThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArg 297
 Db 6103 ACGCAGGATGCTGAAGACCATCACTACAGATAGGGCTTCCCTGACCATCCGC 6162
 QY 298 TyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMet 317
 Db 6163 TACCGTCAGATTGGGCCCTGATTGACCGACAGATCTTCGGCTTCACTGAGGAGGATG 6222
 QY 318 ValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAla 337
 Db 6223 GTCACGCCCGTTTGGTACATACAACTATGACAAACAGCTTCGGGTGACCATGAGGCT 6282
 QY 338 VallIleAsnGluThrProLeuProIleAspIleuTyrArgTyrAspAspValSerGlyLys 357
 Db 6283 GTGATCAACAGACACCCCATGCGCATATGATCTCTATCGCTATGATGATGTGTCAGGCAAG 6342
 QY 358 ThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThr 377
 Db 6343 ACAGAGCAGTTTGGGAAGTTTGGTGTCTTACTATGACATTAACAGATCATCACCACA 6402
 QY 378 AlaValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyr 397
 Db 6403 GCTGTCATGACCCACACCAAGCATTTTGTATGTCATATGGCAGGATGAAGAAGTGCAGTAT 6462
 QY 398 GlnIlePheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgVal 417
 Db 6463 GAGATCTTCCCGTCGCTCATGACTGATGATGACCGCTCCAGTATGATTAACATGGGCGAGTA 6522
 QY 418 ValLysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyr 437
 Db 6523 GTGAGAGAGAGCTGAGTAGGACCCCTACGCCCAATACCATCGCTACTTCTATGAGTAT 6582
 QY 438 AspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSer 457
 Db 6583 GATGCTGACGGCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGC 6642
 QY 458 TyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrPro 477
 Db 6643 TAGACCTCAATGGGAACCTGCACTTACTGAGCCCTGGGACAGTGCACGGCTCACACCA 6702
 QY 478 LeuArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAsp 497
 Db 6703 CTACGGTATGACATCCGACCGCATCCTCGGCTGGGTGACGTGCAATACAAAGATGAT 6762
 QY 498 GluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeu 517
 Db 6763 GAGGATGGCTTCTGAGGACGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGGCTG 6822
 QY 518 LeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeu 537
 Db 6823 CTATCAAGGCTTACAAACCGGGCTGGCAGCTGGAGTGTGAGTACCGCTACGATGGCTG 6882
 QY 538 GlyArgArgVal 541
 Db 6883 GGGCGGGCGCTG 6894

RESULT 5

US-10-342-887-1743
 ; Sequence 1743, Application US/10342887
 ; Publication No. US20040058340A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter S.
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Christopher J.
 ; APPLICANT: van 't Veer, Laura Johanna
 ; APPLICANT: van de Vijver, Marc J.
 ; APPLICANT: Bernards, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-188-999
 ; CURRENT APPLICATION NUMBER: US/10/342.887
 ; CURRENT FILING DATE: 2003-01-15

; PRIOR APPLICATION NUMBER: 60/298,918
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 60/380,710
 ; PRIOR FILING DATE: 2002-05-14
 ; PRIOR APPLICATION NUMBER: 10/172,118
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 1743
 ; LENGTH: 3614
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-342-887-1743

Alignment Scores: 2.88e-257 Length: 3614
 Pred. No.: 2233.00 Matches: 404
 Score: 88.91% Conservative: 77
 Percent Similarity: 74.68% Mismatches: 60
 Best local Similarity: 78.49% Indels: 0
 Query Match: 13 Gaps: 0
 DB:

US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-342-887-1743 (1-3614)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
 Db 272 TACCAGATTGGTTATGACGGCTCCCTCAGATTTATCTACGCCAGTGGCTGGACTCACAC 331
 QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
 Db 332 TACCAAAACAGACCGCAGCTTCTGGCTGSCACCGCTAATCCGACGGTTGCCAAAAGAAC 391
 QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
 Db 392 ATGACTTTCCCTGGCGAGAACGGTCAAAACTTGGTGGAAATGGAGATTCCGAAAAGAGCAA 451
 QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
 Db 452 GCCCAAGGAAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCTT 511
 QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
 Db 512 TCGATTGACTTTGATCGAACACAAAGACAGAAAGATCTATGACGACCAACCGTAATTT 571
 QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
 Db 572 CTACTGAGGATCGCTACGACACGCTGGGCAACCGACTCTCTGGCTGCCAAGCAGCAAG 631
 QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
 Db 632 CTGATGGCGCTCAATGTACCTATTATTCACAGGTCAAATTCGCCAGCATCCAGCGAGGC 691
 QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
 Db 692 ACCACTAGCGAAGAGTAGATTATGACGGACAGGGAGGATCGTGTCTCGGCTCTTTGCT 751
 QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
 Db 752 GATGGTAAAAACATGGAGTTACATATTTAGAAAAGTCCATGGTTCTTCTGCTTCATAGC 811
 QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
 Db 812 CAGCGGAGTACATCTTCGAATACGATATGTGGACCGCTGTCTGCCATCACCATGCC 871
 QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
 Db 872 AGTGTGGCTCGCCACACCATGACAGCATCCGATCCATGGCTACTACCGCAACATATAC 931
 QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
 Db 932 AACCCCGGAAAGCAACGGCTCCATCATCAGGACTACACAGAGAGGGGCTGTTCTA 991
 QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260

```
Db 992 CAACAGCTTTCTGGGTCAAGAGTCGGAGGCTCTTATTCAAATACAGAGGCGAGACTAGG 1051
Qy 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 1052 CTCCTCAGAAATTTATATGATAGCACAGAGTCAGTTTACCTATGATGAACACAGCAGA 1111
Qy 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 1112 GTCCTAAAGACAGTAACCTCCAGAGTCAGTGTGTTTATGACCATGATAGATACAGGCAA 1171
Qy 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 1172 ATTGCTCCCTGATTGACAGGCGAGATTTTCGCTTTAGTGAAGATGGCATGTAATGCA 1231
Qy 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 1232 AGATTGCTATGACTATGATGACCAACAGCTTTTCAGTGACCATGATGAGGCTGTGATCAAT 1291
Qy 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
Db 1292 GAAACGCCACTGCCTATTGATCTGATCTGATTCAGTTTGATGACATTTTCGGCAAGTTGAGCAG 1351
Qy 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 1352 TTTGCAAGATTGGAGTTTATATATGATATTAACAGATCAATTTCTACAGCTGTAATG 1411
Qy 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 1412 ACCTATAGAGAGCACTTTGATGCTATGTCGCGCTATCAAGGAGATTCATATGATATTC 1471
Qy 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 1472 AGGTCGCTCATGTAATGATCAATTCAGTATGATTAACATGGTCCGGTAAACCAAGAGA 1531
Qy 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 1532 GAGATTAATAATAGGCGCTTTGGCAACACACCAATATGCTTATGATATGATGTTGAT 1591
Qy 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
Db 1592 GGACAGCTCCAAACAGTTTACCTCAATGAAGAGATAATGTGGCGGTACAACTACGATCTG 1651
Qy 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 1652 AATGGAACCTCCATTTACTGAACCCAGTAAACAGTCGCGCTCTGACACCCCTTCGCTAT 1711
Qy 481 AspileArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 1712 GACCTCGACAGAGAACTCACTCGACTGGGTGATGTTCAATATCGGTTGGATGAAGATGGT 1771
Qy 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys 520
Db 1772 TTCCTACGTCAAAGGGGCGAGAAATCTTGAATATAGCTTCAAGGGGCTTCTAACTCGA 1831
Qy 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 1832 GTTTACAGTAAGGCGAGTGGCTGCAGCAGTGATCTACCGTTATGACGGCTGGGAAGCGGT 1891
Qy 541 Val 541
Db 1892 GTT 1894
```

RESULT 6

```
US-10-172-118-1743
; Sequence 1743, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
```

```
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1743
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_018104
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1743
```

Alignment Scores:

```
Pred. No.: 2.88e-257 Length: 3614
Score: 2233.00 Matches: 404
Percent Similarity: 88.91% Conservative: 77
Best Local Similarity: 74.68% Mismatches: 60
Query Match: 78.49% Indels: 0
DB: 13 Gaps: 0
```

US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-172-118-1743 (1-3614)

```
Qy 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 272 TACCAGATTGGTTATGACCGCTCCCTCAGAAATTTATCTACGCCAGTCGCTGCATCACAC 331
Qy 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 332 TACCAACACAGACGCGCAGCTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAGAAGAAC 391
Qy 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 392 ATGACTTTGCTGGCGAGAACGGTCAAACTTGGTGAATGGAGATTCCGAAAAGAGCAA 451
Qy 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeu 80
Db 452 GCCCAGGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGCGAGAAACCTCCTT 511
Qy 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 512 TCAGTTGACTTTGATCGAACACAAACACAGAGAAAGATCTATGACGCCACCGTAAATTT 571
Qy 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 572 CTACTGAGGATCCCTTACGACAGCTCTGGGCACCCGACTCTCTGGCTGCCAAGCAGCAAG 631
Qy 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 632 CTGATGGCGCTCAATGTCCATTATTCATCCACAGGTCAAAATTCGCCAGCATCCAGCAGGC 691
Qy 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 692 ACCACTAGCGAGAAAGTAGATTATGACGGACACAGGGAGGATCGTGTCTCGGGTCTTTGCT 751
Qy 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 752 GATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGGTTCTTCTGCTTCATAGC 811
Qy 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 812 CAGCGCGCAGTACATCTTCGATACCATATGTCGGACCGCTGTGCGCATCCATCCGCC 871
Qy 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 872 AGTGTGGCTCGCCACACACCATGACCATCCGATCCATTTGGTGTACTACCGCAACATATAC 931
Qy 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
```


QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5190 TACCAAAACAGAGCCGACCGCTTCTGGCTGGCCACCGCTAATCCGACCGTGGCCAAAGAACA 5249
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluThrArgGlnArgLysGluGln 60
Db 5250 ATGACTTTTGGCTGGCGAAGACCGTCAAAACCTTTGGTGAATGGAGATTCCGAAAGAGCAA 5309
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5310 GCCCAAGGAAAGTCAATGTCTTTGGCCGAGACTCAGGGTTAATGGCAGAAACCTCCTT 5369
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleThrAspAspHisArgLysPhe 100
Db 5370 TCAGTTGACTTTGATCGAACAAACAGACAGAAAGATCTATGACGACCCCGTAAATTT 5429
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuThrSerProSerSerArg 120
Db 5430 CTACTGAGGATCGGCTACGACACCGTCTGGCACCCGACTCTCTGGCTGCCAAGCAGCAAG 5489
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5490 CTGATGCCGCTCATGTACCTATTCATCCAGGTCAAATGGCAGCATCCACGCGAGGC 5549
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5550 ACCACTAGCGAGAAAGTAGATTATGACGACAGGGAGGATCGTCTCGGGTCTTTGCT 5609
QY 161 AspGlyLysThrTrpSerTyrThrTyrIleuGluLysSerMetValLeuLeuLeuHisSer 180
Db 5610 GATGGTAAACATGGAGTTACACATATTTAGAAAGTCCATGGTCTTCTGCTTCATAGC 5669
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuLeuSerValThrMetPro 200
Db 5670 CAGCGCAGTACATCTTCGATACGATATGTGGACCGCTGCTGCCATCACCATCCCC 5729
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5730 AGTGTGGCTGGCACACCATCGACACCATCGATCCATGGCTACTACCGCAACATATAC 5789
QY 221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 5790 AACCCCGGAAAGCAACCGCTCATCATCAGGACTACACGAGGAAGGGCTGCTCTTA 5849
QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 5850 CAAACAGCTTCTTGGGTACAGTCGAGGCTTTTATTCAATACAGAGGAGACTAGG 5909
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 5910 CTCACGAAATTTTATATGATAGACACAGAGTCCAGTTTACCTATGATGAAACAGCAGGA 5969
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 5970 GTCCCTAAAGACAGTAACCTCAGAGTGTGGTTTATTGACCACTATAGATACAGCCAA 6029
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6030 ATTGGTCCCTGATGACAGGCAGATTTCCTGGCTTTAGTGAAGATGGGATGTTAATGCA 6089
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 6090 AGATTGATATAGTATGACAAACAGCTTTCGAGTGACGACGATGCGGGTGTGATCAT 6149
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db 6150 GAAACGCCACTCCCTATTGATCTGATCATGATTTGATGATTCCTGCAAAAGTTGAGCAG 6209
QY 361 PheGlyLysPheGlyValIleTyrTyrIleAsnGlnIleThrThrAlaValMet 380
Db 6210 TTTGGAAGTTTGGAGTTATATATGATATTAACAGATCATTTCTACAGCTGTAATG 6269
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400

Db 6270 ACCTATAGCAAGACACTTTGATGCTCATGCCCGTATCAAGGAGATTCAATATGATATTC 6329
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 6330 AGTCCCTCATGTCTGGATTACAATTCAGTATGATAACATGGGTGGGTAAACCAAGAGA 6389
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 6390 GAGATTAAATAGGCGCTTTGCCACACACCAACCAATATGCTTATGAATATGATTTGAT 6449
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
Db 6450 GGACAGCTCCAAACAGTTTACTCTCAATGAAAAGATAATGTGGCGGTACAACTACGATCTG 6509
QY 461 AsnGlyAsnLeuHisLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 6510 AATGGAAACCTCCATTTACTGAACCCCAAGTAACAGTGGCGTCTGACACCCCTTCGCTAT 6569
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 6570 GACCTCGCAGACAGACATCACTCGACTGGGTGATGTTCAATATCGTTGGATGAAGATGGT 6629
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys 520
Db 6630 TTCTACGTCAAGGGGCGACGAAATCTTTGAATATAGTCCCAAGGGGCTTCTAACTCGA 6689
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 540
Db 6690 GTTTACAGTAAAGCAGTGGCTGGACAGTGATCTACCGTTATGACGCCCTGGGAAGCGT 6749
QY 541 Val 541
Db 6750 GTT 6752
RESULT 8
US-10-038-854-41
; Sequence 41, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spvtek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29

Mon Aug 16 09:01:11 2004

```

; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-41

Alignment Scores:
Pred. No.: 1,11e-256 Length: 8487
Score: 2233.00 Matches: 404
Percent Similarity: 88.91% Conservative: 77
Best Local Similarity: 74.68% Mismatches: 60
Query Match: 78.49% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-038-854-41 (1-8487)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db 5126 TACCAGATTGGTTATGACGGCTCCCTCAGAAATATCTACGCCAGTGGCTGGACTCACAC 5185
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5186 TACCAAAACAGAGCGGCAGCTTCTGGCTGGCAGCGCTAATCCGCGTTCGCCAAAGAAAC 5245
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5246 ATGACTTTGCTGGCGAGAACGGTCAAAACCTTGGTGAATGGAGATTCCGAAAGAGCAA 5305
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5306 GCCCAAGGGAAGTCAATGTCTTTGGCGGCAAGCTCAGGGTTAATGGCAGAAACCTCTT 5365
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 5366 TCAGTTGACTTTGATCGAACACAAACAGACAGAAAGATCTATCAGCAGCACCCTAAATTT 5425
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerArg 120
Db 5426 CTACTGAGGATCGCTACGACACGCTGCGCACCCGACTCTCTGGCTGCCCAAGCAGCAAG 5485
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5486 CTGATGCCCTCAATGTCACCTATTCTATCCAGGTCAATTTGCCAGCATCCAGCGAGGC 5545
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5546 ACCACTAGCGAAGAGTAGATTATGACGACAGCGGAGGATCGGTCTCGGGTCTTTGCT 5605
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGlnLysSerMetValLeuLeuHisSer 180
Db 5606 GATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGGTTCCTTCCTTCATAGC 5665

```

```

181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 5666 CAGCGCAGTACATCTTCGAATATGATATGGACCGCTGTCTGCATACCATGCC 5725
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5726 AGTGTGGCTCGCACACCATCGACACCATCGATTCGATTCGCTACTACCGCAACATATAC 5785
QY 221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 5786 AACCCCGGAAACCAACGCTCCATCATCAGGACTACACAGAGGAGGGCTGCTCTA 5845
QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 5846 CAAACAGCTTCTTGGGTACAAAGTCGAGGGTCTTATTCAAATACAGAGGACGACATAGG 5905
QY 261 LeuAlaGlnThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 5906 CTCTCAGAAATTTTATATGATAGCACAGATCAGTTTACCTATGATGAAACAGCAGGA 5965
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 5966 GTCTAAAGACAGTAAACCTCCAGAGTGTGTTTATTTCACCATTAGATACAGGCAA 6025
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6026 ATTGGTCCCTGATTGACAGGCAGATTTTCGCTTTAGTGAAGATGGGATGTTAATGCA 6085
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 6086 AGATTGTGACTATAGCTATGACACACAGCTTTTCGAGTGACAGCATCGAGGTGTGATCAAT 6145
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db 6146 GAAACGCCACTGCTTATTGATCTGATCAGTTTGTATGATGATTTCTGGCAAGTTGAGCAG 6205
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 6205 TTTGAAAGATTTGGAGTTATATATATGATATTAACAGATCATTTCTACAGCTGTAATG 6265
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 6266 ACCTATACGAGACACTTTGATGCTCATGCGCTATCAAGGAGATTCAATATGAGATATTC 6325
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValLysLys 420
Db 6326 AGGTGCTCATGCTACTGATTAACAATTCAGTATGATAACATGGTGGGTAAACCAAGAGA 6385
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 6386 GAGATTAATAATAGGGCCCTTTGCCAACACCCAAATATGCTTATGAATATGATGTTGAT 6445
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
Db 6446 GGCAGCTCCAAACAGCTTTACCTCAATGAAAGATAATGTGGCGGTACAACTACGATCTG 6505
QY 461 AsnGlyAsnLeuHisLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 6506 AATGGAACCTCCATTTACTGAACCCCAAGTAAACAGTCCGCGCTGTGACACCCCTTCCTAT 6565
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 6566 GACCTCGAGACAGAAATCACTCGCTGGGTGATGTTCAATATCGGTTGGATGAAGATGGT 6625
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys 520
Db 6626 TTCTACGTCAAAGGGGACGAAATCTTTGAATATAGCTCCCAAGGGGCTTCTTAACTCGA 6685
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 540
Db 6686 GTTTTACAGTAAAGCGCATGGCTGACAGTATCTACCGTTATGACCGCTGGGAGGCGGT 6745
541 Val 541
QY

```


341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
 6322 GAAACGCCACTGCTATGTATCTGTATCACTTTGATGACATCTTCTGCAAGTTGACGAG 6381
 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIlelleThrThrAlaValMet 380
 6382 TTTGGAAAGTTTGAGTTATATATATATATATTAATTAACCAACAGATCATTTCTACAGCTGAATG 6441
 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
 6442 ACCATACGAAGCACTTTGATGCTTCATGGCGGTCAAGGAGATTCATTAATGATGATATTC 6501
 401 ArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
 6502 AGGTCGCTCATGTAATGATTAATTCATTAATTCATTAATTCATTAATTCATTAATTCATTAATTC 6561
 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
 6562 GAGATTAATAATPAGGCGCTTTGGCCACACACCAACCAATATGCTTATGAATATGATTTGAT 6621
 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
 6622 GCAGAGCTCCAAACAGTTTACCTCAATGAAAGATATATGGCGGTACAACTACGATCTG 6681
 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
 6682 AATGGAACCTTCATTTACTGAACCCCAAGTAACAGTGGCGTCTGACACCCCTTCGCTAT 6741
 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
 6742 GACCTGCGAGACAGATCACTCGACTGGTGATGCTTCAATATCGTTGGATGAAGATCGT 6801
 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys 520
 6802 TTCCTACGTCAAAAGGGGCACGCAATCTTTCAATATAGCTCCAAAGGGGCTTCTAACTCGA 6861
 521 AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
 6862 GTTTACAGTAAGGACGATGGCTGGACAGATGATCTACCGTTATGACGGCTGGAGGCGGT 6921
 541 Val 541
 6922 GTT 6924
 RESULT 10
 US-10-038-854-35
 ; Sequence 35, Application US/10038854
 ; Publication No. US20040022781A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Spytex, Kimberly A
 ; APPLICANT: Li, Li
 ; APPLICANT: Wolenc, Adam R
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Eissen, Andrew J
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Malyankar, Uriel M
 ; APPLICANT: Shinkets, Richard A
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Spaderna, Steven K
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Gusev, Vladimir Y
 ; APPLICANT: Gangolli, Esha A
 ; APPLICANT: Guo, Xiaojia S
 ; APPLICANT: Shenoy, Suresh G
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Casman, Stacie J
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Burgess, Catherine E
 ; APPLICANT: Edinger, Shlomit R
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: MacDougall, John R
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-230
 ; CURRENT APPLICATION NUMBER: US/10/038,854
 ; CURRENT FILING DATE: 2003-01-22
 ; PRIOR APPLICATION NUMBER: 60/258,928
 ; PRIOR FILING DATE: 2000-12-29
 ; PRIOR APPLICATION NUMBER: 60/259,415
 ; PRIOR FILING DATE: 2001-01-02
 ; PRIOR APPLICATION NUMBER: 60/259,785
 ; PRIOR FILING DATE: 2001-01-04
 ; PRIOR APPLICATION NUMBER: 60/269,814
 ; PRIOR FILING DATE: 2001-02-20
 ; PRIOR APPLICATION NUMBER: 60/279,832
 ; PRIOR FILING DATE: 2001-03-29
 ; PRIOR APPLICATION NUMBER: 60/279,833
 ; PRIOR FILING DATE: 2001-03-29
 ; PRIOR APPLICATION NUMBER: 60/279,863
 ; PRIOR FILING DATE: 2001-03-29
 ; PRIOR APPLICATION NUMBER: 60/283,889
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: 60/284,447
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/286,683
 ; PRIOR FILING DATE: 2001-04-25
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 411
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 35
 ; LENGTH: 8675
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-038-854-35
 Alignment Scores:
 Pred. No.: 1,15e-256 Length: 8675
 Score: 2233.00 Matches: 404
 Percent Similarity: 88.91% Conservative: 77
 Best Local Similarity: 74.68% Mismatches: 60
 Query Match: 78.49% Indels: 0
 DB: 17 Gaps: 0
 US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-038-854-35 (1-8675)
 Qy 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
 Db 5314 TACCAGATTGGTTATGACGGCTCCCTCAGAAATATCTACGCCAGTGGCTTGGACTCACAC 5373
 Qy 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
 Db 5374 TACCAACACAGCGCAGCGTCTTGGCTGGCCACCGCTTAATCCGACGGTTGCCAAAGAAAC 5433
 Qy 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60
 Db 5434 ATGACTTTGCTTGGCGAGAACCGGTCAAACTTGGTGGAAATGGAGATTCCGAAAGAGCAA 5493
 Qy 61 AlaArgGlyGlnValThrValPheGlyArgLeuArgValHisAsnArgAsnLeuLeu 80
 Db 5494 GCCCAAGGAAAGTCAATGCTTTGGCCGCAAGCTCAGGGTTAATGCAGAAACCTTCCT 5553
 Qy 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
 Db 5554 TCAGTTGACTTTGATCGAACACAAAGACAGAAAGATCTATCAGCACACCGCTAAATTT 5613
 Qy 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrPsrProSerSerArg 120
 Db 5614 CTACTGAGGATCGCTACGACAGCTCTGGGCGCACCCGACTCTCTGGCTGCCAAGCAGCAAG 5673
 Qy 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
 Db 5674 CTGATGGCGCGTCAATGTCACTATTTCATCCAGGTCAATATGTCAGCATCCAGCGAGGC 5733

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
 Db 5734 ACCACTAGCGAGAAAGTAGATTATGACGACGAGGAGGAGTCTGCTCGGGTCTTTGCT 5793
 QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
 Db 5794 GATGGTAAACATGAGTATACACATATTTAGAAAAGTCCATGGTCTTCTGCTTCATAGC 5853
 QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
 Db 5854 CACGGCAGTACATCTCGATACGATATGCGGACCGCTGCTCGCCATCACCATGCC 5913
 QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
 Db 5914 AGTGTGGCTCGCCACACCATCGATCCGATCCGATCGGACCGCTGCTCGCCATCACCATGCC 5973
 QY 221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
 Db 5974 AACCCCGGAAAGCAACGCTCCATCATCAGGACTACACGAGGAGGAGGCTGCTTCTA 6033
 QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
 Db 6034 CAACACGCTTCTGGTACAGTCGAGGGCTCTATTCAATACAGAGGAGGAGCTAGG 6093
 QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
 Db 6094 CTCTCAGAAATTTATATGATAGCACAGAGTCAGTTCATCATGATGATGAACAGCAGGA 6153
 QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
 Db 6154 GTCTTAAGACAGTAACCTCCAGAGTGATGTTTATTTGACCAATAGATAGCAGCAA 6213
 QY 301 IleGlyProLeuLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
 Db 6214 ATTGGTCCCTGATGACGACGATTTTCGCTTTAGTAGAGATGGATGATGAATGCA 6273
 QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
 Db 6274 AGATTTGACTATAGCTATGACACAGCTTTCCAGTGACCATGCAGGCTGTGATCAAT 6333
 QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
 Db 6334 GAAACGCACTCGCTATGATCTGTATGATGATGATGATGATGATGATGATGATGATGAT 6393
 QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
 Db 6394 TTTGAAAAGTTTGAGTTAT 6453
 QY 381 ThrHisThrLysHisPheAspAlaTyrArgMetLysGluValGlnTyrGluIlePhe 400
 Db 6454 ACCTATACGAGCACTTTGATGCTCATGCGCGTATCAAGGAGATTCATATATGATATTC 6513
 QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
 Db 6514 AGGTGCTCAGTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6573
 QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
 Db 6574 GAGATTAATAATAGGCGCTTTGCCAACACCAACCAATATGCTTATGATATATGATGAT 6633
 QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
 Db 6634 GGACAGCTCCAAACAGTTTACCTCAATGAAAAGATAATGTGCGCGTACACATGATCTG 6693
 QY 461 AsnGlyAsnLeuHisLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
 Db 6694 AATGAAACCTCCATTTACTGAACCAAGTACAGTGGCGGCTGCTGACACCCCTTCGCTAT 6753
 QY 481 AspileArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
 Db 6754 GACCTGCCGAGACAGAAATCACATGCTGGGTGATGTTCAATATCGGTGATGATGATGAT 6813

QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
 Db 6814 TTCTTACGTCAAAAGGCGCACGGAATCTTTGATATAGTCTCAAGGGCTTCTTAACTCGA 6873
 QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
 Db 6874 GTTTACAGTAAAGGAGTGGCTGGACAGTATCTACGCTTATGACGCTTGGGAAGCGT 6933
 QY 541 Val 541
 Db 6934 GTT 6936
 RESULT 11
 US-09-808-602-74
 ; Sequence 74, Application US/09808602
 ; Patent No. US2002015115A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vernet, Corine A
 ; APPLICANT: Fernandes, Elma
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Herrman, John L
 ; APPLICANT: Majumder, Kumud
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Mezes, Peter S
 ; APPLICANT: MacDougall, John
 ; TITLE OF INVENTION: No, US2002015115A1el Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 15966-697 CIP
 ; CURRENT APPLICATION NUMBER: US/09/808,602
 ; CURRENT FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 09/800,198
 ; PRIOR FILING DATE: 2001-03-05
 ; PRIOR APPLICATION NUMBER: 60/186,596
 ; PRIOR FILING DATE: 2000-03-03
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 74
 ; LENGTH: 8797
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 US-09-808-602-74
 Alignment Scores:
 Pred. No.: 1,19e-245 Length: 8797
 Score: 2141.50 Matches: 382
 Percent Similarity: 87.25% Conservative: 90
 Best Local Similarity: 70.61% Mismatches: 68
 Query Match: 75.27% Indels: 1
 DB: 9 Gaps: 1
 US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-808-602-74 (1-8797)
 QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
 Db 5384 TACCAGCTCTGCAATAATGGAACCTCGGGTATGATGACCAACGGCATGGCTGTGAGC 5443
 QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
 Db 5444 TTCACAGTGAAGCCCGCCGCTCTCGAGGACCAATCACCACCCACCATCGGGCGGTGCAAC 5503
 QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60
 Db 5504 ATCTCTCTGCCCATGGAGATGGCTGAACTCCATCGATGCGCGCTGAGGAAGAACAG 5563
 QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
 Db 5564 ATCAAGCAAAAGTCAACCATCTTTGGGAGGAAGTTCGGGTCCACGGAAGAAATCTCTCG 5623
 QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
 Db 5624 TCCATTGATTAAGCCGAAATATCGGTACGAGGAGATCTACGATGACCAACCGGAAATTC 5683
 QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerSerArg 120

Db 5684 ACCCTGAGGATCATCTATGACCAGTGGCGGCCCTTCTCTGGTCTCCGACGAGTGGG 5743
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5744 CTGGCAGCGTCAATGCTCTCTACTTCTCAATGGGCGCTTGGCCGCGCTTCCAGCGAGGG 5803
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5804 GCCATGAGCGAGGACAGACATTTGACAAGCAGCGCGGATCGTGTCCCGCATGTTCCG 5863
QY 161 AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 5864 GACGGAAAGTCTGGAGTATTCTCTATCTTGCAAGTCCATGCTCTCTCTGTACAGAGC 5923
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 5924 CAAAGTCAGTACATATTGATATGACTCTCTCGATCGCTCCAGCAGTCACTATGCCCC 5983
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5984 AGTGTGCGCGGCACAGCATGTCACGCACACCTCCATTGTTTACATCCGAAACATTTAC 6043
QY 221 GlnProGluGluAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 6044 AACCCACCCGAAAGCAATCATCGGTTCATCTTTGACTACAGTGTGACGCGCGCATCTTA 6103
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 6104 AAGACATCTTCTTGGGCATCTGGCGCCAGGTGTCTACAAGTATGGAATCTTCCAAG 6163
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 6164 TTATCAGACATGATCTTACGACAGCACAGCCGTTCACCTTTGGGTATGACGAGACCCGCT 6223
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 6224 GTCTTGAGATGTCAATCTCAAGTGGGGCTTCTCTGTTACCATCAGTACCGAAAG 6283
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6284 GTTGGGCCCTTGTGCAAGCAGATTTACAGTTCTCTGAGGAGGATGATCAACGCC 6343
QY 321 ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339
Db 6344 AGGTTGATATACCTATACGACAAATAGTTCGCAATTCGCGCATCAAAACCCGCTCAT 6403
QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu 359
Db 6404 AGCGAGACTCCCTCTCTGTTGACCTTACCGCTATGACGATTTCCGCAAGTGGAA 6463
QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379
Db 6464 CACTTCGGCAAGTTGGGGTCTATCTACTACGACATCAACGATCATCAACCATGCGCTC 6523
QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db 6524 ATGACGCTTACCAAGCACTTTGACACCATCGGCGCATCAAGGAGTGCATATGACATG 6583
QY 400 PheArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValValLys 419
Db 6584 TTCCGCTCCCTCATGTACTGTGATGACTGTGCAATATGACATGAGTGGGTGATCAAG 6643
QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db 6644 AGGAACTGAAACTAGGCGCCCTATGCCAAACACCAAAAGTACACCTATGACTATGACGG 6703
QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAsp 459
Db 6704 GACGCCAGCTCCAGAGTGGCGTCAATGACCGGCTTACTCGCGCTATAGTACTATGAC 6763
QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db 6764 CTCAATGGGAACCTGCATCTTCTAAACCCAGAAACAGTGTCTCGCTCATGCCCTTACGC 6823

QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db 6824 TATGACCTCGGTGACCGGATAACACGAGCTAGGGAGCTGACATAAAATCGATGACGAT 6883
QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
Db 6884 GGCATTGTCGACAGAGGTCAGACATCTTTGAATACAACCTCCAAGGCGCTTCTGACG 6943
QY 520 LysAlaTyrAsnArgAlaGlySerTyrPheValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 6944 AGAGCATACAACAAGCCAGCGATGCGAGCGTGCAGTACCGCTATGACGAGTGGCGCG 7003
QY 540 Arg 540
Db 7004 CGG 7006
RESULT 12
US-09-808-602-77
; Sequence 77, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 77
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-808-602-77
Alignment Scores:
Pred. No.: 1,19e-245 Length: 8797
Score: 2141.50 Matches: 382
Percent Similarity: 87.25% Conservative: 90
Best Local Similarity: 70.61% Mismatches: 68
Query Match: 75.27% Indels: 1
DB: 9 Gaps: 1
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-808-602-77 (1-8797)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 5384 TACCAGCTCTCAATATATGGAACCTCGGCTGATGTACGCCAACCGCATCGCTGTGACG 5443
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5444 TTCCACAGTGAAGCCCGGCTCTCGAGGACCATCACCCCATCGGCGGTGCAC 5503
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluThrArgGlnArgLysGluGln 60
Db 5504 ATCTCTCTGCGCATGGAATGGCTTCACTCCATCGATGCGCTCGAGGAAGAAACAG 5563
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgValHisAsnArgAsnLeuLeu 80
Db 5564 ATCAAGGCAAGTCAACATCTTTGGAGGAGTTCGGGTCCACGAGGAATCTCTCTG 5623
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspHisArgLysPhe 100

```
Db 5624 TCCATTGATTATGACCGAATAATATCCGTACCGAGAGAGATCTACGATGACCAACCGGAATTC 5683
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 5684 ACCCTGAGGATCATCTATGACACAGGTGGCGCGCCCTTCCTGTGGCTCCCGGACAGTGGG 5743
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5744 CTGGACCGCTCAATGCTCTCTACTTCTTCAATGGCGGCTTGGCGGCTTCCAGCGAGGG 5803
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5804 GCCATGACGAGAGAGACAGCATTTGACACAGCGCGGATCGTGTCGCCCATGTTCGCC 5863
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
Db 5864 GACGGGAAAGTCTGGAGTTATCTCTTCTTGTGACAAAGTCCATGTCTCTCTGTACAGAGC 5923
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 5924 CAACGTCAGTACATATTTGAATATGACTCCTCCGATCGCTCCAGCAGTCATATGCCCC 5983
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5984 AGTGTCGCCCGGACAGCATGTCCACGACACCTCCATTTGGTTACATCCGAAATTTAC 6043
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 6044 AACCCACCCGAAAGCAATGATCGTCCATCTTTGACTACAGTGATGACGCGCGCATCCTA 6103
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 6104 AAGACATCTTTCTGGGCACTGGCGCGCAGGTGTCTTACAAAGTATGAAATCTCTCCAG 6163
QY 261 LeuAlaGlnThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 6164 TTATCAGAGATAGTCTACGACAGCAGCAGCGTCACCTTTGGGTATGACGAGACCCGGT 6223
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 6224 GTCTGAAGATGTCATCTCCAAAGTGGGGCTTCTCTGTACCATCAGTCAGTACCGAAG 6283
QY 301 IleGlyProLeuLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6284 GTTGGGCGCCCTTGGCAAGACGAGATTACAGGTCTCTGAGGAAGAAATGATCAACGCC 6343
QY 321 ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339
Db 6344 AGGTTTGATTATACCTATACGACAAATAGCTTCGCAATTCGCAATCAACCCGTCATT 6403
QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGlu 359
Db 6404 AGCGAGACTCCCTCTCTGTTGACCTTACCGGTATGACGAGATTTCCGCAAGGTGAA 6463
QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379
Db 6464 CACTTCGGCAAGTTTGGGGTCATCTACTAGACATCAACACGAGATCATCACCTGCCGTC 6523
QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db 6524 ATGACGCTTAGCAAGCACTTTGACACCCATGGGCGCATCAAGGAATGCAATATGAGATG 6583
QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValLys 419
Db 6584 TTCGGTCCCTCATGTACTGATGACATGTGCAATATGACAGTATGGGTAGGGTTCATCAAG 6643
QY 420 LysGlnLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db 6644 AGGGAATGAACCTAGGCGCCCTATGCAACACCAAGTACACCTATGACTATGACGCGG 6703
QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAsp 459
```

```
Db 6704 GACGGCCAGCTCCAGAGTGTGGCGTCAATGATGACCGGCTACCTGGCGCTATAGCTATGAC 6763
QY 460 LeuAsnGlyAsnLeuHisLeuSerProGlyAsnSerAlaArgIleThrProLeuArg 479
Db 6764 CTCAATGGGAACCTGCACCTTCTAAACCCAGGAAACAGTGTCTGCCTCATGSCCTTACGC 6823
QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db 6824 TATGACCTCGTGACCGGATACAGGATAGGGAGCTGTCAGTACAAATTCGATGAGAT 6883
QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
Db 6884 GGCTATTTGGCCAGAGAGGTCAGACATCTTTGAATACAACCTCCAAGGCGCTTCTGACG 6943
QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 6944 AGAGCATACAACAAGCCAGCGGATGAGCGTGCAGTACCGCTATGACGAGTGGGCGCG 7003
QY 540 Arg 540
Db 7004 CGG 7006

RESULT 13
US-09-800-198-62
; Sequence 62, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-62

Alignment Scores:
Pred. No.: 1,19e-245 Length: 8797
Score: 2141.50 Matches: 382
Percent Similarity: 87.25% Conservatives: 90
Best Local Similarity: 70.61% Mismatches: 68
Query Match: 75.27% Indels: 1
DB: 10 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-800-198-62 (1-8797)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 5384 TACCAGCTCTGCAATATGGAACCTCGGGTGTATGATGATGACCAACGCGCTGTCTCAGC 5443
QY 21 LeuGlnThrGlnProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5444 TTCACAGTGAAGCCCAACGCTCTCGGAGGACCATCACCCCGGCGCTGCAAC 5503
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5504 ATCTCTCTGCGCCATGGAGAATGGCTGAACTCCATGAGTGGCGCTTGAGAGGAACAG 5563
QY 61 AlaArgGlyGlnValThrValPheGlyArgLeuArgValHisAsnArgAsnLeuLeu 80
```


5564 ATCAAGGCAAGTACCATCTTTGGGAGGAAGCTTCGGGTCCACGGAAGGAATCTCCTG 5623
Db
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db
5624 TCCATTTGATTATGACCGAATATCCGTACGGAGAGATCTACGATACACCGCGAATTC 5683
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerArg 120
Db
5684 ACCCTGAGGATCATATGACAGAGTGGCGCGCCCTTCCTGCTCCCGAGCAGTGGG 5743
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db
5744 CTGGCAGCGCTCAATGTCTCTTCTTCAATGGCGCTTGGCGCGCTCCAGCAGGG 5803
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db
5804 GCCATGAGGAGAGACAGATTCACAGAGAGCGGATCGTGTCCGCGCATGTTCGCC 5863
QY 161 AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
Db
5864 GACGGGAAGTCTGGAGTTATCTCTATCTTGACAGTCCATGTCTCTCTGTACAGAC 5923
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db
5924 CAACGTCAATATTTGAATGATGACTCTCCGATCGCTCCAGCGAGTCACTATGCC 5983
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrArgAsnIleTyr 220
Db
5984 AGTGTGCGCGGACAGCATGTCCAGCACACTCCATTTGTTATCCGAAACATTTAC 6043
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeu 240
Db
6044 AACCCACCGAAAGCAATGCATCGTCTCTTTGACTACAGTATGACGGCGCATCTTA 6103
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db
6104 AAGACATCTTCTGGGCACTGGCGGCCAGGTGTCTACAGTATGGAATCTCCAA 6163
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db
6164 TTATCAGAGATAGTCTACAGACAGCAGCGGTACCTTTGGTATGACGAGACCCGGT 6223
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db
6224 GTCTCAAGATGGTCAATCTCCAAAGTGGGGCTTCTCTGTATGACATCAGGTACCGAAG 6283
QY 301 IleGlyProLeuLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db
6284 GTTGGGCGGCTGTGGCAAGCAGATTTACAGGTCTCTGAGGAAGGAATGATCAACGCC 6343
QY 321 ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339
Db
6344 AGGTTGATTATACCTATACAGCAATAGTTCGCCATTCGCGATGCAACCCCGTCAAT 6403
QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu 359
Db
6404 AGCGAGACTCCCTTCTCTGTGTGACCTACCGCTATGACGATTTCCGGCAAGGTGAA 6463
QY 360 GlnPheGlyLysPheGlyValIleTyrThrAspIleAsnGlnIleIleThrThrAlaVal 379
Db
6464 CACTTCGGCAAGTTGGGGTCTATCTTACGACATCAACAGATCATCACCACCTCCGCTC 6523
QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db
6524 ATGACGCTTAGCAAGCACTTTGACACCCATGGCGCATCAAGGAAGTCAATATGAGATG 6583
QY 400 PheArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValValLys 419
Db
6584 TTCCGGTCCCTCATGTACTGTGATGACTGTGCAATATGACAGTATGGGTAGGTCATCAAG 6643
QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db
6644 AGGNACTGAACACTAGGCGCCCTATGCCAACACCAACAAAGTACCTATGACTATGACGG 6703

QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAsp 459
Db
6704 GACGGCCAGCTCCAGAGTGGCGCTCAATGACCGGCTTACCTGGCTATAGCTATGAC 6763
QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db
6764 CTCATGGGAACCTGCACCTTCTAAACCCAGGAAACAGTGTCTCGCTCATGCCCTTACGC 6823
QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db
6824 TATGACCTCCGTGACCGGATACACAGGTAGGGAGCTGACGACCAAAATCGATGACGAT 6883
QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
Db
6884 GGCTATTTGTGCCAGAGGGTTCAGACATCTTTTGAATACACTCAAGGGCTTCTGACG 6943
QY 520 LysAlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArg 539
Db
6944 AGAGCATACAAAGCCGAGCGGATGGAGCTGCGAGTACCGCTATGACGGAGTGGCGCG 7003
QY 540 Arg 540
Db
7004 CGG 7006

RESULT 14

US-09-800-198-65
; Sequence 65, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-65

Alignment Scores:
Pred. No.: 1,19e-245 Length: 8797
Score: 2141.50 Matches: 382
Percent Similarity: 87.25% Conservative: 90
Best Local Similarity: 70.61% Mismatches: 68
Query Match: 75.27% Indels: 1
DB: 10 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-800-198-65 (1-8797)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db
5384 TACAGCTCTCCATATATGGAACTCTGGGGTGTATGACGCCAACCGCATGGCTGTGACG 5443
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db
5444 TTCCACAGTAGGCCACCGTCTCGAGGACCATCACCCCAACCATCGCGGCTGCAAC 5503
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGluArgLysGluGln 60
Db
5504 ATCTCTTGCCCATGGAGATGGGCTGAACCTCCAGTGGCGCTGAGGAAGGAACAG 5563


```
QY 61 AlaArgGlyClnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
DB 5564 ATCAAGGGCAAGTCAACCATTTTGGAGGAAGCTTCGGGTCCACGGAAGAAATCTCTCG 5623
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
DB 5624 TCCATTGATTATGACCGAATATCCGTACGGAGAGATCTACGATGACCCACCGGAATTC 5683
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrPsrProSerArg 120
DB 5684 ACCCTGAGGATCATCTATGACCAAGTGGCGGCCCTTCTGTGCTCCCGAGCAGTGGG 5743
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyTyrIleAlaGlyIleGlnArgGly 140
DB 5744 CTGGCAGCCCTCAATGTCCTCACTCTTCAATGGGGCGCTTGGCGGCCCTCCAGCGAGGG 5803
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
DB 5804 GCCATGAGCGAGAGGACAGACATGTGACAAGCGCGGATCGTGTCCCGCATGTTCCGCC 5863
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
DB 5864 GACGGGAAGCTCTGGAGTTATCTCTATCTTGACAAGTCCATGCTCTCTGTCTACAGAGC 5923
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro 200
DB 5924 CAACGTCAGTACATATTTGAATATGACTCTCCGATCGCTCCAGCGCATCATATGCC 5983
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
DB 5984 AGTGTCCCGCGGCACAGCATGTCCACGCACACCTCCATTTGTTATCCGAAACATTTAC 6043
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
DB 6044 AAACCCACCCGAAAGCAATGCATCGGTCTTTGACTACAGTGCATGACGCGCGCATCCTA 6103
QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
DB 6104 AAGACATCTTTCTTGGGCATCGGCGCGCGAGTGTCTACAAAGTATGGAAACTCTCCAAG 6163
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
DB 6164 TTATCAGAGATAGTCTACGACAGCAGCGGTACCTTTGGGTATGACGAGACCCCGGT 6223
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
DB 6224 GTCTGGAAGATGCTCAATCTCCAAAGTGGGGGCTTCTCTGTACCATCAGGTACCGAAG 6283
QY 301 IleGlyProIleLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
DB 6284 GTTGGGCCCCCTTGGCAAGACGAGATTACAGGTCTCTGAGGAAGGAATGATCAAGCC 6343
QY 321 ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339
DB 6344 AGTTTGTATTATACCTATACGACAAATAGCTTCGGCATTCGCACATCAAAACCGTCAAT 6403
QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGlu 359
DB 6404 AGCGAGACTCCCTCTCTGTGTGACCTACCGGTATGACGAGATTTCCGCAAGGTGAA 6463
QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379
DB 6464 CACTTCGGCAAGTTTGGGGTCTATCTACGACATCAACACGATCATCACCTCGCCGTC 6523
QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
DB 6524 ATGACGTTAGCAAGCATTTTGACACCCATGGGGCGCATCAAGGAGTGCATATGAGATG 6583
QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValLys 419
DB 6584 TTCGGTCCCTCATGTACTGATGACTGTGCAATATGACAGTATGGGTGAGGTGCATCAAG 6643
```

```
QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
DB 6644 AGGGAAGTGAACCTAGGCGCCCTATGCCCACACCAAAAGTACACCTATGACTATGACGGG 6703
QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAsp 459
DB 6704 GACGGCCAGCTCAGAGTGTGGCGGTCAATGACCGGCTACCTGGCGCTATAGCTATGAC 6763
QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
DB 6764 CTCATGGGAACCTGACACCTTCTAAACCCAGGAAACAGTCTCGCCTCATGCCCTTACGC 6823
QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
DB 6824 TATGACCTCGTGACCGGATACAGCTAGGGGAGCTGCAGTACAAATTCGATGACGAT 6883
QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
DB 6884 GGCTATTGTGCCAGAGAGGGTCAGACATCTTTGAATACAACTCCAAGGCGCTTCTGACG 6943
QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539
DB 6944 AGAGCATACAACAAGCGCCAGCGGATGGAGGTGCAGTACCGCTATGACGAGTGGCGCGC 7003
QY 540 Arg 540
DB 7004 CGG 7006

RESULT 15
US-09-808-602-79
; Sequence 79, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 79
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-808-602-79

Alignment Scores:
Pred. No.: 2,55e-245 Length: 8409
Score: 2138.50 Matches: 384
Percent Similarity: 86.69% Conservative: 85
Best Local Similarity: 70.98% Mismatches: 71
Query Match: 75.17% Indels: 1
DB: 9 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-808-602-79 (1-8409)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
DB 5386 TACCAGCTCTGTAATATGTTACTTTGAGAGTGTATGATGCCAATGCCATGATATTAGC 5445
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
```

Db 5446 TTTACACGCAACCTCATCTGCTGGGACAGTAACCTCCACCATAGGACGATGTAAT 5505
Qy 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60
Db 5506 ATTTCTCTACCAATGGAGATGGTTTGAACTCAATTTGAATGGCGTCTGAGGAAGAACAG 5565
Qy 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5566 ATTAAGGCAAGAGTCTGCTGTTGGAAGAAGCTCAGGGTTTCATGAAGAAATTTGCTG 5625
Qy 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 5626 TCCATGTATACGACCGGAATATACGCCACAGAAAAATCTACGATCATCACCGCAAGTTC 5685
Qy 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerSerArg 120
Db 5686 ACCCTGAGGATATTTACATCAGCTGGACGGCCCTTCTCTGGCTGCCACGACGGC 5745
Qy 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5746 CTGGCTGCGCTCAACGTCTTATTTCTTCAACGGCGCGCTGGCTGGCTTCAGCGCGA 5805
Qy 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5806 GCCATGCGCAAGGACACATCCACAAAGCAAGCAGCATCATATCGCGCATGTTTCCA 5865
Qy 161 AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 5866 GATGGGAAGTTTGGAGTTTACACCTACCTAGAAAAATCCATGGTACTACTGCTTCAGAGC 5925
Qy 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 5926 CAGCGCGAGTACATCTTTGAGTATGATCTTTCAGACGGCGCTCCATGCTGTTACTATGCT 5985
Qy 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5986 AGTGTGTGCGCATAGATGCAATCCACAGCTCTGTGGCTCATTTAGGAATATTTAT 6045
Qy 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 6046 AATCTCTCTGAAGCAACGCATCATGATTTTGTATTACAGTATGATGGGAGGATTTTG 6105
Qy 241 HisThrPheThrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 6106 AAAACATCATTTTATAGTACTGGTCGACAAAGTCTTTTACAGTATGGAAGCTATCCAAA 6165
Qy 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 6166 TTATCTGAAATGTTTATGACAGTACTCGCGTTACTTTTGGATATGATGAAACTACAGT 6225
Qy 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 6226 GTCCTAAAATGGTGAATTTGCAAGTGGAGGATTTCTTGTACANTCGCTATCGTAAA 6285
Qy 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6286 ATTGCCCTCTTGTGACAAAACAAATCTACAGATTTCTCTGAAGAAGTATGTTCAATGCA 6345
Qy 321 ArgPheAspTyrAsnTyr--AspAsnSerPheArgValThrSerMetGlnAlaValIle 339
Db 6346 AGTTTGTATATACATATCACACAAATAGTTTTCGATTTGCAAGCATCAAACTCATATA 6405
Qy 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu 359
Db 6406 AGTGAGACTCCCTCTCCAGTTGATCTTTACCGTTATGATGATTTCTGGCAAAAGTTGAG 6465
Qy 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleThrThrAlaVal 379
Db 6466 CATTTTGGCAATTTGGAGTTATTTATATGATATAAAATCAAAATATTACTACAGAGTT 6525
Qy 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db 6526 ATGACACTGAGTAAGCACTTTGATACCCACGACGCAATTAAGAAGTTCAATATGAGATG 6585

Qy 400 PheArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValVallys 419
Db 6586 TTCCGATCCCTGATGTACTGGATGACTGTGCANTATGACAGCATGGGAAGAGTAATAA 6645
Qy 420 LysGluLeuLysValGlyProTyrTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db 6646 AGAGAACTGAAACTTTGGGCGGTATGCCAACACAAACCAAGTATACCTATATGATGGA 6705
Qy 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAsp 459
Db 6706 GATGGCAATTGCAAAAGCGTAGCAATAATGATAGCCCTTACCTGGCGTTTACAGTTATGAC 6765
Qy 460 LeuAsnGlyAsnLeuHisLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db 6766 CTGAATGGAATCTTCACCTCTGATCTCTGGAACAGTGTTCGATTCGATGCCCTGCGC 6825
Qy 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db 6826 TACGACCTCAGAGACAGGATTCACGCTTAGGTGACATACCGTACAAAATCGATGATGAC 6885
Qy 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
Db 6886 GGAATCTCTGTGTCAACGAGGCTCAGATGTATTTGAGTACAAATTCCAAAGGACTTTTACA 6945
Qy 520 LysAlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 6946 AGAGCTTACAACAAAGCAAAATGGATGGAAACGTTTCAGTACCGTTACGACGGACTTGGCGA 7005
Qy 540 Arg 540
Db 7006 AGG 7008

Search completed: August 14, 2004, 19:21:47
Job time : 1215.27 secs

| | | | |
|----|------|---|------|
| Qy | 27 | LeuLeuAlaGlyThrValAsnProThrValGlyArgAsnValThrLeuProIleAsp | 46 |
| Db | 2070 | -----GCAGCCGCGCGTCCATCATCGGCGAG----- | 2111 |
| Qy | 47 | Asn---GlyLeuAsnLeuValGluTrpArgGlnArg----- | 61 |
| Db | 2112 | GGACTCGCGTACGTCTGGAAACGAGACGCGAGGTTACCGCATCACGCCAGCGCAG | 2171 |
| Qy | 62 | ArgGlyGlnValThr-----ValPheGlyArgArgLeuArg | 73 |
| Db | 2172 | CGAGGAAGTCTATGCTCCACGACGACAACGCGCGCTGGTACGCCAGGTGCATCCGACGG | 2231 |
| Qy | 74 | ValHisAsnArgAsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIle | 93 |
| Db | 2232 | AGCGAAACCCCTCAGGCACTACGACGAGGAGGCGCAACTGGTCGCCGAAACGCGATCGGTT | 2291 |
| Qy | 94 | TyrAspAspHisArgLysPheThrLeuArgIle-LeuTyrAspGlnAlaGlyArgProSe | 113 |
| Db | 2292 | GGGAACGATCAC-----CGAGTACCGCTACGACGAGCGCGACGCTTGGA | 2336 |
| Qy | 113 | rLeuTrpSerProSerSerArgLeuAsnGlyValAsnValThrTyrSerPro---GlyG | 132 |
| Db | 2337 | GGCACTGCTCCCGGCC-----GAAGCGAGCCAACTGCTATAGCTACTTCGACGG | 2387 |
| Qy | 132 | YTrpIleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGlnTyrAspGlnAlaG | 152 |
| Db | 2388 | CTTCTGGCGAGCGTGGCCGCGGAGCGCGAGTGAAGTACGAGCGCAACGCTCAGG | 2447 |
| Qy | 152 | YArgIleThrSerArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGlu | 172 |
| Db | 2448 | CGACATACCCCGCAGACCGATCCGGAAGCAAGTCACCCATACGCTACGCTACGACCCG | 2507 |
| Qy | 172 | sSer---MetValLeuLeuLeuHis----- | 179 |
| Db | 2508 | CGGTGCGTGGTGGAAATCACCCTATCGCGAGCGGAGCTGCACGTTGACCTGGAAATCC | 2567 |
| Qy | 180 | -----SerGlnArgGlnTyr----- | 184 |
| Db | 2568 | GCTCGGCCAGTTGTAGAAGAAACAGCTTCGCGACGGTAGCTGGCGCGCTATCGCTACGA | 2627 |
| Qy | 185 | -----IlePheGluPhe----- | 188 |
| Db | 2628 | CACACTGGGTAGACAGATCACCCGCCAGGACGAAAGCGGCCCATCACCGGTTCCAAATG | 2687 |
| Qy | 188 | eAspLysAsnAspArgLeuSerValThrMetProAsnValAlaArgGlnThrLeuG | 208 |
| Db | 2688 | GGATGCAGCAGCGCGCTGACGAGATCACCTCGCGGTGGCGCCAGCCGATG----- | 2742 |
| Qy | 208 | uThrIleArgSerValGlyTyrTyrArgAsnIleTyr-----GlnProPr | 223 |
| Db | 2743 | -----TACCGCTACAACGCTTAGCGCAAGGTCAACGAGCTACCGAGTAGTCGA | 2783 |
| Qy | 223 | oGluLysAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPh | 243 |
| Db | 2784 | CGAACAGCGCGGATGACCCGCTACGATATCTTGACGACTGCACCTGGTGGTACCGCGG | 2843 |
| Qy | 243 | eTyrLeuGlyThrGlyArgArgValIleTyrLysTyr----- | 255 |
| Db | 2844 | GATCAACCCGAGCGGCACTGAGCTCGGCTACCGCTACGAAACCGCGGCTACTGCTCAG | 2903 |
| Qy | 256 | -----GlyLysLe----- | 258 |
| Db | 2904 | CGAGATCGAGAACGAAACGCGGCGAACGCTATCGCGTGTATACACGCAACGCGCTGAT | 2963 |
| Qy | 258 | uSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluTh | 278 |
| Db | 2964 | CAGCCAG-----GAAACCGGCTTCGAGCGCGCGCACCGCTATCGTACGACCTCAA | 3017 |
| Qy | 278 | rAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTy | 298 |
| Db | 3018 | AGGCCAGGCTTCGGGAGAAACCGAGTACGCGGACGAGCGAGCGAGCACTGCCACACCTA | 3077 |
| Qy | 298 | rArgGln-----IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGlu | 316 |

| | | | | | |
|----|------|---------------|--|--------------------|------|
| Db | 3078 | CCAGCGCGA | CAGTACCGGGCGCTGCTGGGAAGACACTG | -----CCGACGG | 3125 |
| Qy | 316 | yMetValAsnAla | AargPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetG1 | 336 | |
| Db | 3126 | C----- | AATCGGTCGACTACCGCTACGATACGCTCGGACGCGTGGTCGCC | ----- | 3171 |
| Qy | 336 | nAlaValIleAsn | GluThrProLeuProIleAspLeuTyrArgTyrAspValSerG1 | 356 | |
| Db | 3172 | ----- | GTCGACGACGACCTGGCGCTG---GCTCAAGATACGAC | ----- | 3210 |
| Qy | 356 | yLysThrGluGln | PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleTh | 376 | |
| Db | 3210 | ----- | ----- | ----- | 3210 |
| Qy | 376 | rThrAlaValMet | ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValG1 | 396 | |
| Db | 3211 | ----- | -----CTACG | ----- | 3215 |
| Qy | 396 | nTyrGluIlePhe | ArgSerLeuMetTyrTrpMetThrValGln-----TyrAspAsnMe | 414 | |
| Db | 3216 | GGACCGCGTGT | TCGCGAGCACGAGGCTGGCCACCTTCAC | TACGCCCTACGACGCCCT | 3275 |
| Qy | 414 | tGlyArgValVal | LysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSe | 434 | |
| Db | 3276 | CGGCCAGTTGAT | CCACTGCCACTGCCGACGGC-----AACCGCGTCGA | 3320 | |
| Qy | 434 | rTyrGluTyrAsp | AlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTr | 454 | |
| Db | 3321 | CTACCGCTAC | CAGACCGCGCACCTCTCGGCAATCGATCTCAACGGCCAGTCCCTGAC | 3380 | |
| Qy | 454 | pArg----- | ----- | ----- | 455 |
| Db | 3381 | CCGSCACCACTT | CGGTAGCGTCCGAAACGCCAGCCAGCGGCGAACTGCTCAGCCA | 3440 | |
| Qy | 456 | -TyrSerTyrAsp | LeuAsnGlyAsnLeu----- | ----- | 465 |
| Db | 3441 | ATACCACTAC | GATGAGCAGGCGCCCTGCTGGCCACCAGGTCTCAGTCAGCGGACGCGCCA | 3500 | |
| Qy | 465 | sLeuLeu----- | -----SerProGlyAsnSerAlaArgLeuThrProLe | 478 | |
| Db | 3501 | TCTCTACGAG | CGCCAGTACGCTATGAGCTAGCGCAACTGCGCCGCTCAGGACAG | 3560 | |
| Qy | 478 | uArgTyrAspIle | Arg-----AspArgIleThrArgLeuGlyAspValGln-- | 493 | |
| Db | 3561 | CCGCAAGGCA | TTCGAGCTTCCACTACGACCCGCTGAGCCGCTGCTCGGCGTACGCGG | 3620 | |
| Qy | 494 | ----- | -----TyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyAs | 508 | |
| Db | 3621 | TGAACCCCGA | AAAGCTTCGTGACGACCCGCGCAACTGCTCGCCAGGCGGCGCA | 3680 | |
| Qy | 508 | pile----- | ----- | ----- | 509 |
| Db | 3681 | ATTCGACGCG | CAGACAGATGGAGTTCCGGGAATCGCTGCTGACGAGCGGCGACCGCCA | 3740 | |
| Qy | 510 | -PheGluTyrAsn | SerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlagly----- | 526 | |
| Db | 3741 | TTTCGACTAC | CAGCCCACTCGTTCGGAACGCGCGGCAACCGGCGAGAGCT | 3800 | |
| Qy | 527 | ----- | ----- | ----- | 527 |
| Db | 3801 | GGTCACCGAA | TACGCTACGACTGCCAACCCGCTGATCGGCGTACGCTGCGGACGG | 3860 | |
| Qy | 527 | rTrpSerValArg | TyrArgTyrAspGlyLeuGlyArgVal | 541 | |
| Db | 3861 | GCGCAGAGT | CTCTACCGCTACGACGCGCTTCGGCGGACGATC | 3903 | |

RESULT 2

; Sequence 304, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 304
 ; LENGTH: 4998
 ; TYPE: DNA
 ; ORGANISM: Proteus mirabilis
 ; US-09-543-681A-304

Alignment Scores:
 Pred. No.: 6,42e-15 Length: 4998
 Score: 208.00 Matches: 140
 Percent Similarity: 31.40% Conservative: 87
 Best Local Similarity: 19.36% Mismatches: 224
 Query Match: 7.31% Indels: 272
 DB: 4 Gaps: 37

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-543-681A-304 (1-4998)

QY 2 TyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAlaLeu 21
 DB 2200 TACATTGACGCTGAGGCT-----GGATTTTCAGTTATATAT 2235
 QY 22 GlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsnVal 41
 DB 2236 TATAATGATGATAATCTGGTGACTCGCACTATCGATCCA---TTATGGCGTGAACGTTAT 2292
 QY 42 ThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgGlyGluGlnAla 61
 DB 2293 ACG-----GAAATGGGAACAGCGGAAAGATAGCA 2322
 QY 62 Arg-----GlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArg 77
 DB 2323 GAAATCAATGATGTTGGTCAGCGCACAGATATGCT-----TATCACTATATAT 2370
 QY 78 AsnLeuLeuSerLeu-----Asp 83
 DB 2371 GGACTGCTTGGCTATATCTATTACCTGATGTAAGCCATTCTATGATTATATATGAC 2430
 QY 84 PheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPheThrLeuArg 103
 DB 2431 TATGTCACCTTACCATTTCATCGCGCTTTGGCGACGAA-----TGGCAG 2478
 QY 104 IleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArgLeuAsnGly 123
 DB 2479 CTAAGCTATGATGAAATGGC-----AAT 2502
 QY 124 ValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGlyIleMetSer 143
 DB 2503 TTAACCATAGTGACAGATCCGCAAGG-----CGTCAG 2535
 QY 144 GluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAlaAspGlyLys 163
 DB 2536 CAGGCTCTATGATATAGCAACATGGTGAATGCTCAAGACCAATTACGCCCAATGGTCG 2595
 QY 164 ThrTrpSerTyrThrTyrLeuGluLysSerMetValLeu----- 176
 DB 2596 CAATGGCAATACCACTATAACCTGCACATCACTGATTATAAAACCAACCAATCCTTATCAA 2655
 QY 177 -----LeuLeuHis----- 179
 DB 2656 CACAGCACAGATATCATAGTATGAATGGGCGCTTTACTGCACCTACACTGATGCGTTA 2715
 QY 180 -----SerGlnArgGlnTyrIlePheGluPhe-----AspLysAsnAspArgLeuSer 195
 DB 2716 AATCACACCACCGCTATCAATATACAGCACTGAGCATGCCAGTACCAACGCGAGTCTCAGC 2775
 QY 196 SerValThrMetProAsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyr 215

DB 2776 AAGATTTTATTACCCGAT-----GGGGTC 2799
 QY 216 TyrArgAsnIleTyrGlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGlu 235
 DB 2800 GAGCAACACATAGACTATGACAGTGAAGCCCGGAGTTATCGGTGACGGAT----- 2850
 QY 236 AspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyr 255
 DB 2851 -----GGTGAAGGAAACACCCCGTTATCGTTAT 2880
 QY 256 GlyLysLeuSerLysLeuAlaGluThrLeuTyr-----AspThrThrLysValSerPheThr 274
 DB 2881 GGGCCTTTTGATGTCTTACTGCAATGATACGCCCTTCACGCGCAGTGAATTCGCTTTGAA 2940
 QY 275 TyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGlu----- 290
 DB 2941 TATGATTCACTGACACGACTTAAAGAAAGTAGTGAATGCGAACGGGAGGTTTATCTTTAC 3000
 QY 291 -----GlyPheThrCysThr 295
 DB 3001 GAACGGGATAAAGCGGACAAATCATTCGTGAAGTTGATTTTACGGGCGGTGAGATTGCT 3060
 QY 296 IleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIle----- 309
 DB 3061 TATCGCTATGACGACTAGGACGTGCATTGCCACTCGCTATCCGATAACCATGAACGTG 3120
 QY 310 -----PheArgPheThrGluGluGlyMetVal----- 318
 DB 3121 CGTTGGCGCTAATGAAGTAGGCTTAGTGGTTGAACAAGCGAATGTTTGAAGATGAG 3180
 QY 319 -----AsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArg 331
 DB 3181 CAAGAAGACCGCTGTTTATCGACGCTCAATATAGCTATAAC----- 3222
 QY 332 ValThrSerMetGlnAlaValIleAsnGluThr-----ProLeuProIleAspLeuTyrArg 350
 DB 3223 -----GCAGTCACAAATTGATAAAGCGCAATCTCTGACTCGTGGTGGTGAATTGAG 3276
 QY 351 TyrAspAspValSerGlyLys-----ThrGluGlnPheGlyLysPheGlyValIleTyr 368
 DB 3277 TATGACGAC---CAAGCGCGTTTATGCGAGTGAACGTATTAATGAACAAGAGATTGTT--- 3330
 QY 369 TyrAspIleAsnGlnIleThrThrAlaValMetThrHisThrLysHis----- 385
 DB 3331 CATCAATGGAATGAAGCTGATAATACGCTGCGATTGACTCGTTTGGTGAGCGGAGTTG 3390
 QY 386 -----PheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePheArgSer 402
 DB 3391 CATATGCTTTTGGGCGCACTCGGGAGTTAACTTCCTGCAAGTCAATCAGCATGCACCA 3450
 QY 403 LeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeu 422
 DB 3451 CTTCAATTTC-----AGCTACAACGCGGTAGGACAAGAGTATTACGCCGTAGT 3498
 QY 423 LysValGly----- 425
 DB 3499 CGCGCGGATTGTCAATTCCAGTCACTATACGGCTACCGGTTTACTGGCTCATCAACGA 3558
 QY 426 -----Pro 426
 DB 3559 GCAGGGCGAGCCGCAACAGCTTTTACAATCGATACAAGCCCATCCCAACAACCGCCA 3618
 QY 427 TyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla----- 439
 DB 3619 TTTTGTACGGATGTCACCGCAGTTTCAATATGATCGCGCTTATAATGTGTGTGGTATT 3678
 QY 440 -----AspGlyGlnLeuGlnThr-----ValSerIleAsnAspLys----- 451
 DB 3679 GAAGATGATCGCTGCGCACACACGCTTATCACTATAATGCAATGACCAATCACTGAG 3738
 QY 452 -----ProLeuTrp-----ArgTyrSerTyrAspLeuAsnGly 462

| | | | |
|----|------|---|------|
| Db | 1417 | CAGTTAACGTCAGCCACCGCGCTGGAATTCGGCCGGGAATATGATGAATTC | 1476 |
| Qy | 256 | GlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyr | 275 |
| Db | 1477 | GGCCGCTGG-----ATTTCAGGAACATGCCCTCAGCGCGATATCATCCCGCTACCTGTTAT | 1530 |
| Qy | 276 | AspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThr | 295 |
| Db | 1531 | GATATCCACACAGTACGATTACCTCGCAACGGAAGATCCACCGGCAGCCGGAAAAACC | 1590 |
| Qy | 296 | IleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGlu | 315 |
| Db | 1591 | ATGACGTGGAGCCGTTACGGT-----CAGTTGTCTGAGCTTCACCGACTGT | 1635 |
| Qy | 316 | GlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMet | 335 |
| Db | 1636 | TCCGGTTATGTAACCCGTTATGACCATGAC-----CGCTTCGGGCAGATG | 1680 |
| Qy | 336 | GlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArg-----TyrAsp----- | 352 |
| Db | 1681 | ACGCGGCTGCACCGCGAAGAGGCTG-----AGTCAGTACCGCGCATACGACGCGCT | 1734 |
| Qy | 353 | -----AspValSerGlyLysThrGluGlnPheGlyLysPheGly | 365 |
| Db | 1735 | GGACAGTTAATTCGGTGAAAGACACGCGAGGCCATGAACCGCGTAT---GAATACAAC | 1791 |
| Qy | 366 | ValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHisThrLysHis | 385 |
| Db | 1792 | ---ATCGCGCGGTGACCTGACCGCGCTCATTCGCCGCGACGACGAGAAACGGACACAG | 1848 |
| Qy | 386 | PheAspAlaTyrGlyArg---MetLysGluValGlnTyrGluIlePheArgSerLeuMet | 404 |
| Db | 1849 | TACGATCGCTGGGGAAAGCCGCTCCGTACACGACGCGCGGGCTAACCGCCAGTATG--- | 1905 |
| Qy | 405 | TyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysVal | 424 |
| Db | 1906 | -----GAATACAGATGCTCGCGACGGGTCTATCCGCTGACCCAGTGAATAAC | 1950 |
| Qy | 425 | GlyProTyrAlaAsnThrThrArgTyrSer-----TyrGluTyrAsp | 438 |
| Db | 1951 | GGCAGCCAC---ACCACCTCCGTTACCATGTACTGTGACCGCGCTGATACAGAAACCGCG | 2007 |
| Qy | 439 | AlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrTyrSerTyr | 458 |
| Db | 2008 | TTTGACGCGCGACACAG-----CGTTATCACCAC | 2037 |
| Qy | 459 | AspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeu | 478 |
| Db | 2038 | GACCTTGACCGGCAAA-----CTTATCCGACGAGGATGAGGCTGTGTCACCCACTGG | 2091 |
| Qy | 479 | ArgTyrAspIleArgAspArgIleThrArg-----LeuGlyAspVal-----Gln | 493 |
| Db | 2092 | CATTATGCAAGACGACCGCTCAGCACCGCCACCGTGAAGGCTGAACCCGACGAGCGG | 2151 |
| Qy | 494 | TyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsn | 513 |
| Db | 2152 | TGGCAGTATGACGAACGTTGGCTGGGTGACAGAC----- | 2184 |
| Qy | 514 | SerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArg | 533 |
| Db | 2185 | -----ATCAGCATATCATCGCAAGGCGCACCGGTGGCGGTGCATTACAGG | 2229 |
| Qy | 534 | TyrAspGlyLeuGlyArg | 539 |
| Db | 2230 | TATGATCAGAAAGCGCG | 2247 |

RESULT 5

RESULTS
US-09-252-991A-14461

US-09-252-991A-14461
: Sequence 14461. Application US/09252991A

; sequence 14461, Apr
; Patent No. 6551795

Patent No. 6551795
; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

QY 243 -----PheTyrLeuGlyThrGly 248
Db 1576 CTGACCGGCGCGCTACGACCGCTGGGGCAACTGATCGCGCTGGAAATGGCGCGACGC 1635
QY 249 ArgArgValIleTyrLysTyr-----GlyLysLeuSerLysLeuAlaGlu----- 263
Db 1636 AGCGCCCTGAGCTATGAAATACGACGCGCTTGGCGCGACGCCGATATCGCGGATGCCGAG 1695
QY 264 -----ThrLeuTyrAsp----- 267
Db 1696 GGGCAGCGCACCGCTGTCTCTGGGGGACGCGGATCTGCTGGCGGGGTTCCTCCGACGCC 1755
QY 268 ---ThrThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsn 286
Db 1756 GCGCGCGCGGANTGCTTACCTGCGACGAGAA---GCCGCGAGGCTGTGCTGCCCTGACC 1812
QY 287 LeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAsp 306
Db 1813 AACGAGAAC---GGCTCCAGCGCGAGTTCCGCTAGCACCTGTGGACCGTCTGGTGGAA 1869
QY 307 -----ArgGlnIlePheArgPheThrGluGluGlyMetValAsn 319
Db 1870 GAGACCGGCTTCGACGGTCTGCCCAACGCTATCGCTACACGCGCGCTGACGAACGTATC 1929
QY 320 AlaArgPheAsp-----TyrAsnTyrAspAsnSerPheArgVal 332
Db 1930 GCCCGCGAGATGCGCGTCCGAGACCACTATGCTACGCGCGGATGGCGCGCTG 1989
QY 333 ThrSerMetGlnAlaValIleAsnGlu---ThrProLeuProIleAspLeuTyrArg--- 350
Db 1990 GCGAGCATCCGGTTCGCGCACCGGACGACGCGCGCTGTGCGAGCGTACCGCTGG 2049
QY 351 -----TyrGly----- 390
Db 2230 CCGTGGCTGGTGAACCTATGATGGCCCGGGTCACTGACGCGCGCTGGTGGTGGCTC 2289
QY 391 -----ArgMetLysGluValGlnTyrGluIlePheArg--- 401
Db 2290 GAGCTGGCTTCGAGCGCGATGCCCTGACCGTGGGTGGCTGGCGCGCGCGCGCGAC 2349
QY 402 -----SerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgVal 418
Db 2350 GGGCAGGATGACGGCTGTGTTCCACGAGGCGCCAGCAGCGCGCTGGAGCGCTGCAG 2409
QY 419 LysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAsp 438
Db 2410 CGCAGCGCTGCTGCTGGCGCGGGTTCGACTGGCAGCGC---GGCTATCGCTAGCAC 2466
QY 439 AlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyr 458
Db 2467 GGACTTGGACAACTGGTGGCGATCGACGACCAACGATATCGCTGGTGGCTAGCAATAC 2526
QY 459 AspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeu 478
Db 2527 GACCTGCGCGCTGCGCTG---CTGCCAGCGCGCGCTGGCGCGCGCGGCGGACCTAC 2583
QY 479 ArgTyrAsp----- 481
Db 2584 CGATACGACGCGCGCGCAACCGCTCTGGAGGGCGCTGCGCGAGTACGCGCGCGAGCGCT 2643

QY 482 -----IleArg 483
Db 2644 CGCCAGGCATTCGCCGAGAACGAAGTGTACCGCTCCGGTTTCAGTCCGTCGAGGTGGCG 2703
QY 484 -----AspArgIleThrArgLeuGly 490
Db 2704 GCGAACCGCGCGCGGCGGCGCGCGCTGGGCGCGCAACGCGCTGGAGAGGATCGCC 2763
QY 491 AspValGlnTyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePhe 510
Db 2764 GCGAACCGCTATCGCTTCGATCGCTGGCGCAACCTGCTGAGCGTATCGCGCGCGCG 2823
QY 511 Glu-----TyrAsnSerAlaGlyLeuLeuIle-----LysAlaTyr 522
Db 2824 GAGCGCTGCGCTGCGCTTACGACGCGCGCGCGCTGCTCATCTGACGCGCGATAC 2883
QY 523 AsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArgVal 541
Db 2884 GCGGACGCGACGCGCTGGAGCGCGCTACCGCTACGACGCGCTGTGCGCGCGGATC 2940
RESULT 6
US-09-328-352-1416
; Sequence 1416, Application US/09328352
; Patent No 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1416
; LENGTH: 4791
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1416
Alignment Scores:
Pred. No.: 3,96e-13 Length: 4791
Score: 193.50 Matches: 128
Percent Similarity: 30.43% Conservative: 89
Best Local Similarity: 17.95% Mismatches: 211
Query Match: 6.80% Indels: 285
DB: Gaps: 29
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-328-352-1416 (1-4791)
QY 1 TyrTyrIleGlyAlaAspGly---SerLeuArgLeuLeuLeuAlaAsnGlyMetGluVal 19
Db 1837 TATTACTTCGATCTTGATGGTGTACTTATCGAACTCGTTTAGCGGATGGCGTGAGTCA 1896
QY 20 AlaLeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArg 39
Db 1897 TGGTATAGCCGTGAT---GGGAAAAAG 1920
QY 40 AsnValThrLeuProIleAsp----- 46
Db 1921 CGCATTACTGACAAATCGATTTTGATGGCGAGAGACTCAACAAGATACAAATGATCAA 1980
QY 47 -----AsnGly----- 48
Db 1981 GATCAACTGTGTTAAATTTGTTTCAGCCAAACGGTGGGATAATTCGTTTGCATATAAAG 2040
QY 49 ---LeuAsnLeuValGlu-----TrpArgGlnArgLysGlu 59
Db 2041 CAGGTAACTTAGTTGAAATTAAGAGCCAGAGGTAGTAGTTTGGAAACCGAATACGAT 2100
QY 60 GlnAlaArg-----GlyGlnValThrValPheGlyArg 70
Db 2101 GAAATAGAAATGTAGTAAAGAAATCAATCCATTAGGCGATATCACACAGTAC----- 2154
QY 71 ArgLeuArgValHisAsnArgAsnLeuLeuSerLeuAspPheAspArgValThrArgThr 90

| | | | | |
|----|------|--|--|------|
| Db | 2155 | ---- | AAATATAACAATGATAATCAACTGGTTGAAGTCATTGATGCAAAAAGGTGGTGT | 2208 |
| Qy | 91 | GluLysIleTyrAspAspHisArgLysPheThrLeuArgIleLeuTyr---- | AspGlnAla | 109 |
| Db | 2209 | AAAGAAATTT----- | CAGTATATGAATTTGGGCGAGATGATTTT | 2262 |
| Qy | 110 | GlyArgProSerLeuTip----- | SerProSerSer | 119 |
| Db | 2263 | GGTAGAGTAGTACATGGGAATATGATGAAGATGGTGCACGTGACAGCTGAACAACTGCA | 2322 | |
| Qy | 120 | ArgLeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArg | 139 | |
| Db | 2323 | ATAATAAAGATTGTACAGTATTTTACAGTACCAAGCCGAGATAAAGAGACACTGCAA | 2382 | |
| Qy | 140 | GlyIle----- | MetSerGluArgMetGluTyrAspGlnAlaGlyArgIle | 154 |
| Db | 2383 | AGTATTATTATTCCTGATCGGTGTTAAAGATACTTTGACATGATGAGAGAGCGCTTA | 2442 | |
| Qy | 155 | ThrSerArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMet | 174 | |
| Db | 2443 | CTTAACATACCGATACCAAGGTTTAGTAAACGAATATAAGTACAAATCAAGTTGGCTTA | 2502 | |
| Qy | 175 | Val----- | 175 | |
| Db | 2503 | TTAGACGCGTATTGATGCCAATGACATAGTGTAGCTTATCAATGGGCAAAACAAGGT | 2562 | |
| Qy | 176 | ----- | LeuLeuLeuHisSerGlnArgGlnTyrIlePhe | 186 |
| Db | 2563 | CGCATTCMAAAGCTAATCAACAGATCAACGCGAATATTTGTTGGATACAAATCCGAT | 2622 | |
| Qy | 187 | ----- | GluPheAspLys | 190 |
| Db | 2623 | GGATATTGTGATCGTGAACAGCGCTTTGTATGGAGAGGAAACACATAGTTACAAATGA | 2682 | |
| Qy | 191 | AsnAspArgLeuSerSerValThrMetProAsnValAlaArgGln----- | 205 | |
| Db | 2683 | AATGGTCGATTATTTCAATCCGTCGACCAATATATACTGACCAATTTGACTATTATGCA | 2742 | |
| Qy | 205 | ----- | 205 | |
| Db | 2743 | GATGGCAAAATTCCTCGAAAAGTTTATCCCACTTACATACAGGCGCAGAAACACAGAG | 2802 | |
| Qy | 206 | ----- | ThrLeuGluThrIleArgSerVal | 213 |
| Db | 2803 | CAATTTGACTATTAACCTCAATAGCCAACTGAGCCGAGCAGTAATGAGTTAGCCAAAT | 2862 | |
| Qy | 214 | GlyTyrTyrArgAsnIle----- | TyrGlnProPro | 223 |
| Db | 2863 | GATTTATCGCAATGCATCTGGGCGAGTTGGTTAGAGAGCATCAGCAATTTATAAAATCCA | 2922 | |
| Qy | 224 | Glu----- | GlyAsnAlaSer | 228 |
| Db | 2923 | GAACTGAAACCGCTCACAGCAGTATTACATATGAAATATGATGACGCTTGGAAAT | 2976 | |
| Qy | 229 | ValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGly | 248 | |
| Db | 2977 | TTAATTAACAAATTCGACCTGATGGACATACATCTGAATCTTTTATGTTGTTGAGGG | 3036 | |
| Qy | 249 | ArgArgValIleTyrLysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThr | 268 | |
| Db | 3037 | CAT----- | ATTTATGCAATTTGTT | 3066 |
| Qy | 269 | ThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGln | 288 | |
| Db | 3067 | CAGGTTGTTTCATTTTCAGCGTGATGATTG----- | CATAGAGAAACACACGATTACTA | 3120 |
| Qy | 289 | AsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGln | 308 | |
| Db | 3121 | GCAATGGGCTCATGCAACCAACATATATGATGTGGGG----- | TTATGTGTTCCACAG | 3177 |
| Qy | 309 | IlePheArgPheThrGluGluGly----- | MetValAsnAlaArgPheAspTyr | 324 |

```

; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 226:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5519
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 226:
US-09-453-702B-226

Alignment Scores:
Pred. No.: 7,76e-13 Length: 5519
Score: 192.00 Matches: 133
Percent Similarity: 34.70% Conservative: 78
Best Local Similarity: 21.88% Mismatches: 227
Query Match: 6.75% Indels: 170
DB: 4 Gaps: 29

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-453-702B-226 (1-5519)

QY 4 GlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAlaLeuGlnThr 23
Db 822 GGTGCGGGCGTCATCTCGCGCTGTACTGACACGCGAGCGCGCGGCGGAGAAAGCC 881
QY 24 GluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsnValThrLeu 43
Db 882 CGGACGAGCGGCATTTCCGCGGGGAGGAAACGTCGCTTTTCCTGAT-----ACCTCG 935
QY 44 Pro-----IleAspAsnGlyLeuAsnLeu-----ValGluTrpArgGln 56
Db 936 CCGGTTACCGAATATGCGCGGAGCAACGCGATCCGTCGTCGCGGTGGCTGACG 995
QY 57 ArgLysGluGlnAlaArgGlyGlnValThrValPheGlyArgArgLeuValHisAsn 76
Db 996 CACGACCGGAA-----TACCCG 1013
QY 77 ArgAsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAsp 96
Db 1014 GAGAAATTTACCTCGCGCGCGCTGGTGCGC-----TATGCTGG 1052
QY 97 HisArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSer 116
Db 1053 ACGCGCGCGGCGAATCGCGGTGGTGTATGACCGTAGTGGCAACAGAGTGCGC-----1106
QY 117 ProSerSerArgLeuAsnGlyValAsnValThrTyrSerPro-----GlyGlyTyr 133
Db 1107 -----ACCTTACTTACGATGATAAATACCGGCGCGGATG 1142
QY 134 IleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArg 153
Db 1143 GTGGCGACCGCTCACACGCGCGCGCGGAAATCCGTTACCGTTACGACAGCGCGGCGG 1202
QY 154 IleThrSerArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSer 173
Db 1203 GTGACAGAAAGCTAAACCCCGGAGGCTTAAGCTACACGTATCAGTAT---GAGAAAGAC 1259
QY 174 MetValLeuLeu-----LeuHisSerGlnArgGln---183
Db 1260 CGCATCACCATCAGGACAGCGCTGNAACCGCGGTGAAGTCTTCGACACGCGAGGGAAGGC 1319
QY 184 -----TyrIlePheGluPhe 188
;
1320 GGGCTGAACCGGGTGGTGAAGAAAGGAAACACGCGGACGCGAGCGTCAAGAGTCAAGTTT 1379
189 AspLysAsnAspArgLeuSerValThrMetProAsnValAlaArgGlnThrLeuGlu 208
1380 GACGCGGTGGCAGGCTCAGGGCACAGAG-----GATCGCGCAGGCGAGGACACAGAA 1433
209 ThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAlaSer 228
1434 -----TACAGCCCGGATGTGGTGACGGGCTC 1460
229 ValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPheTyr-----244
1461 ATCAGCGGATCACCACGCGGATGCGAGGCGATCGCGCTTTTACTATTAACACACACAGC 1520
245 -----LeuGlyThrGly-----ArgArgValIleTyrLysTyr 255
1521 CAGTTAAACGTACGCCACCGGGCTGACGGGTGGAATATACCGCGGAAATATGATGATGG 1580
256 GlyLysLeuSerLysLeuAlaGlnThrLeuTyrAspThrThrLysValSerPheThrTyr 275
1581 GGCCTCTG-----ATTGAGGAAACTGCCCTGACGCGGATATCACCGCTACCGTTAT 1634
276 AspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThr 295
1635 GATAATCCACACAGTACTTACCTCGCAACGGAAGATGCCACGCGAGCGGAAACCC 1694
296 IleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGlu 315
1695 ATGACGTGGAGCGGTTACGGT-----CAGTTGCTGAGCTTACCGACGCT 1739
316 GlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMet 335
1740 TCCGGTTATGTAACCCGTTATGACCATGAC-----CGTTCCGGGAGATG 1784
336 GlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSer 355
1785 ACGGGGTGACCGCGAGGAAGGCTG-----AGTCAGTACCGCGCATACACAGCGGT 1838
356 GlyLys-----ThrGluGlnPheGlyLys 363
1839 GGACAGTTAATTGCGTGAAAGACACGCGAGGCGCCATGAACGCGGTATGAATACACGCC 1898
364 PheGlyValIleTyrTyrAspIleAsnGlnIleThrThrAlaValMetThrHisThr 383
1899 GCCGGT-----GACCTGACCAACCGCTCATTTGCCCGGACGCGGACGAGAAACGGG 1946
384 LysHisPheAspAlaTyrGlyArg---MetLysGluValGlnTyrGluIlePheArgSer 402
1947 ACACAGTACGATGCGTGGGAAAGCCATCTGTACCACGAGGCGGCTGACGCGCAGT 2006
403 LeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysGluLeu 422
2007 ATG-----GAATACGATGCTCGCGGACGGGTGTCATCCGCTGACCACT 2048
423 LysValGlyProTyrAlaAsnThrThrArgTyrSer-----TyrGlu 436
2049 GAAACCGCACCCAC-----ACCACCTTCGTTACGATGCTGACCGCGCTGATACAGAA 2105
437 TyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyr 456
2106 ACCGGCTTTCAGCGCGGACACAG-----CGTTAT 2135
457 SerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThr 476
2136 CACCACGACCTGACCGGCAAA-----CTTATCCGAGGAGGATGAGGGGTGCTGACC 2189
477 ProLeuArgTyrAspIleArgAspArgIleThrArg-----LeuGlyAspVal---492
2190 CACTGGCACTATGACGAAGACGACCGGCTCACGACCGGACCGGCTGAAGGGTGAACCGCA 2249
493 ---GlnTyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGlu 511
2250 GAGCGTGGCAGTATGACGACGCGGCTGGCTGACAGAC-----2288

```


QY 446 ValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeuAsnGlyValAsnLeuHis 465
 Db 27556 CAGCGGGTTTGTATGCGCGGACGACGTTTACCAATGACCTGACCGGAAAA----- 27609
 QY 466 LeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArgAspArg 485
 Db 27610 CTCACGACAGTGAACGACGAGGGGCTTGTACCCCTCTGGCAGTACGATCGGTGCGACCGC 27669
 QY 486 IleThrArgLeu-----GlyAsp-----ValGlnTyrLysMetAspGluAspGly 500
 Db 27670 ATCAGCACCAGCGGTGAACGCGGACCGGACGAGCAGTGGCAGTATGATGAGCACCGG 27729
 QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
 Db 27730 TGGCTA-----ACCACCTCAGCCATACCGTGAAGGCCACCGGGTG----- 27771
 QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539
 Db 27772 -----TCGGTCCACTACGGCTATGACGATAAAGGCCGC 27804
 RESULT 9
 US-09-328-352-1419
 ; Sequence 1419, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 1419
 ; LENGTH: 2415
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-1419

Alignment Scores:

Pred. No.: 1,27e-12 Length: 2415
 Score: 185.50 Matches: 142
 Percent Similarity: 33.79% Conservative: 78
 Best Local Similarity: 21.81% Mismatches: 232
 Query Match: 6.52% Indels: 199
 DB: 4 Gaps: 33

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-328-352-1419 (1-2415)

QY 33 AsnProThrValGlyLysArgAsnValThrLeuProIleAsp-----AsnGly 48
 Db 220 AATTATGTTATAGTGATCAACATGTGGTGTACCAATGAAGTTGATGTTGGAAATGTT 279
 QY 49 LeuAsnLeuValGluTrpArgGlnArgLysGluGlnAlaArgGlyGlnValThrValPhe 68
 Db 280 AAGAAATTTACTCAATTCGTGGTTACAAAAGGGGTATTCCAACTCAAGTTAAACTTCA 339
 QY 69 GlyArgArgValHisAsnArgAsnLeuLeuSerLeuAspPheAspArgValThr 88
 Db 340 AATGGGGCGACAGAAACA-----AATATTGTAGAT---GACTTTGGAATATAACT 387
 QY 89 ArgThrGluLysIleTyrAspAspHisArgLysPheThrLeuArgIleLeuTyrAspGln 108
 Db 388 CAA-----CATACTGATGCATGGGTATTTTCAAGAAACAATATGATGAT 435
 QY 109 AlaGlyArgProSerLeuTrpSerProSerArgLeuAsnGlyValAsnValThrTyr 128
 Db 436 GCTGCGCGATTATATATAGATACCAATTTGTTGGTTAAATTTACTCTACATTCACITAT 495
 QY 129 -----SerProGlyGlyTyrIleAlaGlyIleGlnArg 139
 Db 496 GATGGTTTAACTGTTTCTAGGGTTGTAACTGGTGGCGGCAACTGAGCCGAATAGAAAA 555

QY 140 Gly-----IleMetSerGlu----- 144
 Db 556 TATAATGGTGATGGTTTACTGATTTTCGAGTGAAGGATAAGATATCAAAATAATCTATTATT 615
 QY 145 ---ArgMetGluTyrAspGlnAlaGlyArgIleThrSerArg----- 157
 Db 616 AACTCTAATAAGTATGATGATTTGGAAACCTGATATTTAAATCCCAATCTCTGGTTTGTAGT 675
 QY 158 IlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeu 177
 Db 676 GCTATAACAAGTGAACACAGCAGTTATGATGCTTTTGACCGCCCAATTACTGTC--- 732
 QY 178 LeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerVal 197
 Db 733 -----AATGATAATGGAAGTGTGTG 753
 QY 198 Thr-----MetProAsnValAlaArgGlnThr-----LeuGluThrIleArgSer 212
 Db 754 ACATATTGCTATCAGTCTCTGTGGAGGAAAACTGTGCAATTTGTACAAACAACAGATAGC 813
 QY 213 ValGlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAlaSerValIleGlnAsp 232
 Db 814 TTTGGT-----ACAAACAGAACTCTAATCTATTAGCAGCAGGAGAT 852
 QY 233 PheThrGluAspGlyHisLeuLeuHisThrPheTyrLeu-----GlyThrGlyArgArg 250
 Db 853 TTTAGTGTCTGAT-----TTAAAAACACTGGTTGCACGTAAGGTACAGATGGATCA 903
 QY 251 ValIle----- 252
 Db 904 GTCCTTCAGACACAACTGAATTTGAAAACGCACTTTTAAACCAAAAGTCTCTGTATGC 963
 QY 253 -----TyrLysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyr 266
 Db 964 GGTAGTTCTACACAAAGTTATACATATACTCTAATAACAACCTTTAGCA---ACAGAAAA 1020
 QY 267 AspThrThr-----LysValSerPheThrTyrAspGluThrAlaGlyMetLeuLys 283
 Db 1021 GATAATAGTATTAGTGGCGCAAAAGACATTTAAATATGATGATCT---GGCCGAATTACT 1077
 QY 284 ThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyPro 303
 Db 1078 TCANTTACCCATCCAGATAGTTCAGTAGAA---ACGATCAATATTTTCAATTAAGGAC 1134
 QY 304 LeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAlaArgPheAsp 323
 Db 1135 CTCATCGGTCCTCCGACATGCGGAGAGTTCAGACA-----ACT 1173
 QY 324 TyrAsnTyrAspAsnSerPheArgVal---ThrSerMetGlnAlaValIleAsnGluThr 342
 Db 1174 TATAGTTATTCTCTGTGGCGGCTTTAAAAACAACAGCTAATGCAAAATATTAGTGAAGCA 1233
 QY 343 ProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGlnPheGly 362
 Db 1234 -----TTTGTCTGTACATATGCGCGCTGTATAGCTATACACAAATAATCAAT 1284
 QY 363 -----LysPheGlyValIleTyrTyrAspIleAsn 372
 Db 1285 GCCAATGATACCAATAATTCATATGCTGTAAGATATGCG-----TATAATCAATTAAC 1338
 QY 373 GlnIleIleThr-----ThrAlaValMetThrHisThrLysHisPheAspAla 388
 Db 1339 CAGTTTACTCAATTCATTAATTCCTAAATGGTAAATCTGTAATTTATCTAATCAGAAATGCT 1398
 QY 389 TyrGlyArgMetLysGluValGlnTyrGluIlePheArgSerLeuMetTyr----- 405
 Db 1399 TTAGGGGAAGTAAACATCAATACCA---AATGTAATACAAAGTTTGAATTAATGACCGG 1455
 QY 406 -----TrpMetThrValGln----- 410
 Db 1456 CAACAACCTTACAACTGTGCAAGCCATACCGCATACTTTTGTGTCATATATCTTATAATGAT 1515
 QY 410 ----- 410

Db 1516 AGCGGGTTCGTAATAATATCTCTGCTACAGTCTCGAAAGTGTCTCTCAATGTTGAT 1575
QY 411 -----TyrAspAsnMetGlyArgValValLysLysGluLeuLysValGlyProTyrAla 428
Db 1576 TATGGTTACGATAAATTGAACGAGTTAATAAGCTTTCGACAAATGTGATCAGTTTAT 1635
QY 429 AsnThrThr----- 431
Db 1636 AAGCTGACCAATTGACCGTTATGGTACAGGTTTGATGAGTACTGTTGAGTTAGATCAAGCA 1695
QY 432 ArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLys 451
Db 1696 CGTTATCAATATAGTTATATAATATGAT-----GACATTACTAAAGTTAATATCACTTCACAAA 1752
QY 452 -----ProLeuTyrArgTyrSerTyrAspLeuAsnGlyAsnLeuHis 465
Db 1753 AGCTGACAGTAGCAGCAGCTACTATACATAATACTAT-----GTAAATAACACTTCTCGA 1809
QY 466 LeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArgAspArg 485
Db 1810 TTAGCAAGCGTAAGTGGCTCAACT-----TATAGCTTTACTTATGATGCGATGGGTAAT 1863
QY 486 IleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGlyPheLeuArgGlnArg 505
Db 1864 GTAAACAATGATGCTGGCTACTCTAACTTATGATAATTATAGTCGCCTCAATAAAGAT 1923
QY 506 GlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLysAlaTyrAsnArgAla 525
Db 1924 GGTAAATGAGACCTATCTTTATATATGCTGATGAGCTAAGATCGGTGCTGTAAGGGATGAT 1983
QY 526 GlySerTrpSerValArgTyrArgTyrAspGly 536
Db 1984 GGTTTAACAGATTATTAATGATCTTGATGGT 2016

RESULT 10
US-09-453-702B-90/c
; Sequence 90, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166

; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4453
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 90:
US-09-453-702B-90

Alignment Scores: Length: 4453
Pred. No.: 3,55e-12 Matches: 116
Score: 185.50 Conservativeness: 68
Percent Similarity: 34.01% Mismatches: 220
Best Local Similarity: 21.44% Indels: 137
Query Match: 6.52% Gaps: 21
DB: 4

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-453-702B-90 (1-4453)

QY 49 LeuAsnLeuValGluTrpArgGlnArgLysGluGlnAlaArgGlyGlnValThrValPhe 68
Db 1467 CTGCTCTACACCGCAGCGCGCGGCGGAGAGCCCGTAAACAG----- 1420
QY 69 GlyArgGluArgValHisAsnArgAsnLeuLeuSerLeuAspPhe----- 84
Db 1419 -----CACACCGCTTCTTTATCTTCCCTGCACACCCCGCCCTCTT 1378
QY 85 -----AspArgValThr 88
Db 1377 TCAGACTCAGCGTTCCCGACACACTGCCCGGTACCGAATACGGTCCGACAGAGTATC 1318
QY 89 ArgThrGluLysIleTyrAspAspHisArgLysPhe----- 100
Db 1317 CGCTTTTCGGGTGTGGCTGACGACGCGCCCGCATACCCGGAAGCCCTGCCGGTGC 1258
QY 101 ---ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSer 119
Db 1257 CCACCTGCGCGGTACAGCTATACGGAAGCGGTGAACCTGTGCGGTATATGACCGCAGC 1198
QY 120 ArgLeuAsnGlyValAsnValThrTyrSer-----ProGlyGlyTyrIleAlaGly 136
Db 1197 AATACGCGAGTGGCGCTTTCAGTATAACGCGCAGCATCCGGCGCGGATGTTGGCCAC 1138
QY 137 IleGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyIleThrSer 156
Db 1137 CGTTACGCGGAAGCCGAGATCGCTACCGCTACGACGATACCGCGCGGTGTGGAG 1078
QY 157 ArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeu 176
Db 1077 CAGCTGAACCCGCGAGCGCTGAGTTACCGCTACCATGATGAGCAGGACCGCATCCCGTC 1018
QY 177 LeuLeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSer 196
Db 1017 ACGACAGCTGAACCGCGGTGAGTCTGCATACAGAGCGCGCGCGGCTGAACCGG 958
QY 197 ValThrMetProAsnValAla----- 203
Db 957 GTGGTGAAGAAAGAACTGCGCGCAGCGGTACGCACACAGCGGTATGACGCGCAGGA 898
QY 204 ArgGlnThrLeuGluThrIle-----ArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 897 AGGCTACGCGCAGCAGCGCGCGCGGACGCGGACAGAGTACCGTCTGAATGTGGTG 838
QY 221 Gln-----ProProGluGlyAsnAlaSerValIleGlnAspPhe 233
Db 837 TCGGCGATATCAGGACATCACACACCGAGCGCGGGGAGACG-----AAATTTTAC 784
QY 234 ThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGly-----ArgArg 250
Db 783 TATAACGACGGAACACGCTGACGCGCGTGTGTGTCCCGGACGCGGTGGAGAGCGCGCG 724
QY 251 ValIleTyrLysTyrGlyLysLeuSerLeuAlaGluThrLeuTyrAspThrThrLys 270

```

Db 723 GCATATGATGAACCGGCGAGGCTG-----GATCGGAGCATCGCGCTGTGGGACGTC 670
QY 271 valSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGlu 290
Db 669 ATCCGGTATGCTTATATGATATCCGACAGTGAATTACCGGCCACGACAACAGATGCGACG 610
QY 291 GlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePhe 310
Db 609 GCGACACCGCGAGTACGTGGAGCCCTACGGG-----CAGTTGCTG 565
QY 311 ArgPheThrGluGluGlyMetValAsnAlaArgPheAspTyrAsn----- 325
Db 564 GCGTTCACTGACTGCTCGGGCTACACAGACCGCTTATGAATACGCGCTTCGCCAGATG 505
QY 326 -----TyrAspAsnSerPheArg 331
Db 504 ACGCGGTTACCGCGGAGGAGGTATACCGCTTACCGCGCTATGACAAACGTTGGCGG 445
QY 332 ValThrSerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyr 351
Db 444 TTAACCTCGTGAAA----- 430
QY 352 AspAspValSerGlyLysThrGlu-----GlnPheGlyLysPheGlyValIleTyrTyr 369
Db 429 ---GACGCACAGGCGGTGAAACGCGGTATGATACAAACCGCGCAGG----- 385
QY 370 AspIleAsnGlnIleIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyr 389
Db 384 GACCTGACTGCGGTATACCCCGGACCGCAACCGGAGGAGACACAGTACGATCGGTG 325
QY 390 GlyArg---MetLysGluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThr 408
Db 324 GGAAAGGGGTCAGACACACAGCGCGGGCTGACCGCGCAGTATG----- 280
QY 409 ValGlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysValGlyProTyrAla 428
Db 279 ---GAGTATGACGCGGACGAGCGTGTACCACCGTGTACCAACGAG----- 238
QY 429 AsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIle 448
Db 237 AACGCGCCGCGAGTGAATTTACCTTACGATGTGTGCGACCGCTGACGAGCAGCGCGC 178
QY 449 AsnAspLysProLeuTyrArgTyrSerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSer 468
Db 177 TTTGATGCGCAGACACACCGTTACCGTACGCGCCACAGG-----CAGCTATCCGC 124
QY 469 ProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArg 488
Db 123 AGCGAGGATGAGGGCAGGTCAACGAGTGTACTACGATGAAGCAGACCGCATCAGGCAC 64
QY 489 Leu-----GlyAsp-----ValGlnTyrLysMetAspGluAspGlyPheLeuArg 503
Db 63 CGGACGGTGAACGCGGACCGCGCAGCAGTGGCAGTATGATGAGCAGCGGTGGTTACAG 4
QY 504 Gln 504
Db 3 AAA 1

```

RESULT 11

```

US-09-252-991A-11853
; Sequence 11853, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

```

```

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11853
; LENGTH: 2997
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11853

Alignment Scores:
Pred. No.: 2,82e-12 Length: 2997
Score: 184.00 Matches: 130
Percent Similarity: 37.14% Conservative: 88
Best Local Similarity: 22.13% Mismatches: 203
Query Match: 6.47% Indels: 167
DB: 4 Gaps: 29

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-252-991A-11853 (1-2997)

QY 56 GlnArgLysGluGlnAlaArgGlyGlnValThrValPheGlyArg----- 71
Db 531 CAACGCGCTCACCG-----CAGCCTCGCTACACCGCGCTCGACGGTGGTGGC 581
QY 72 LeuArgValHisAsnArg-AsnLeuLeuSerLeuAspPheAsp----- 85
Db 582 CTCCTGTCAGCACCGCGGAGCAGCACCGCGCTTCGACTATGACGAGTGGCCAGATCAC 641
QY 86 -ArgValThrArgThrGluLysIleTyrAspAspHisArgLysPheThrLeuArgIleLe 105
Db 642 CCGCGTCAACCGTGGCGACGCGAGTTGG-----CTGAGCTACGA 680
QY 105 uTyrAspGlnAlaGlyArgProSerLeu-Trp---SerProSerSerArgLeuAsnGlyV 124
Db 681 ATACGACACGACGCGCGC-----CTGGTGGCGATCGCAACACCTCGCGGAACGCT 734
QY 124 alAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArg----- 139
Db 735 CGAATACGACTCGACACCAAGGCA---ATCGACC-GCCAGCGCATCAAGGACGCA 790
QY 140 -----GlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerA 157
Db 791 CGCGCAGCGCTGTCGCGCAGCAGCAATGGGCTAGCAGAGTGGCGCGCTGCTCCGTG 850
QY 157 rgIlePheAlaAspGlyLysThrTrpSerTyrThrTyr-----L 170
Db 851 CGGTGCGGCGCGCGCGCAGACACGCGACTTCGCTACGACTCAACACACAATCGGTG 910
QY 170 euGluLysSerMetValLeuLeuLeuHisSerGlnArgGluTyrIlePheGluPheAspL 190
Db 911 GCGAAACCAACCGCGCGCAGTTCGCCACACAGCCAG-----GCCTTCGACG 955
QY 190 ysAsnAspArgLeuSerSerValThrMetProAsnValAlaArgGlnThrLeuGluThrI 210
Db 956 CCCTCGACCGCGCTGTCGCGCAGACGATCC----- 987
QY 210 leArgSerValGlyTyrTyrArgAsnIleTyrGlnProGluGlyAsnAlaSerValI 230
Db 988 -----CTCGCGGCAAGACCGCGCTCGCTACGACGCGCCAGCAACCTCACCGAG 1039
QY 230 leGlnAspPheThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGlyArg 250
Db 1040 TCAAGGAC-----CCGCGCGCGCTCA 1060
QY 250 rgValIleTyrLysTyr-----GlyLysLeuSerLysLeuAlaGluThrLeuTyrA 267
Db 1061 CCACCGCTACGAATACGACGCGCTCGCAACCTGACCGACTCGTCAGCCCG-----G 1114
QY 267 spThrThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnL 287
Db 1115 ACAGCGGCAACACACCTTCGACGACGACCGCGCGCAACAGTCACTCGCGCGCACCG 1174
QY 287 euGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspA 307
Db 1175 CTCGC-----GGTGGGTCAACGAGTATCGTACGACGCGCTCAATCGCTGTCGAGC 1228

```

```
QY 307 rgGlnIlePheArgPheThrGluGluGlyMet---ValAsnAlaArgPheAspTyrAsnT 326
Db 1229 GCGCG-----TCGCGAGAGCCCGAGCGCTCGAGTACAGTACGCTACGCTACCG 1282
QY 326 yrAsnAsnSerPheArgValThrSerMetGlnAlaValIleAsnGlnThrProLeuProI 346
Db 1283 CCGAGCGCAACAGGCGATCGCGCGCTGGCGCCATC-----GAAGGTGCGCGGACA 1336
QY 346 leAspLeuTyrArgTyrAspValSerGlyLysThrGluGlnPheGlyLysPheGlyV 366
Db 1337 GCTGTGTGTACCGTACGAGCGCGCAACCTGCTCGAGCAGTACGCGAGC----- 1389
QY 366 alIleTyrTyrAspIleAsnGlnIleThrThrAlaValMetThrHisThrLysHisP 386
Db 1390 --ATCGCGCTCGACAGCAGACCTGCTCGACCGCGTG-----ACCTACCGCT 1435
QY 386 heAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePheArgSerLeuMetTyr- 405
Db 1436 ACAGCGGCGCAACCACTGCTGAGATCGGTACCCCTCGGCTCGCCATCGGTACC 1495
QY 405 ----- 405
Db 1496 CGCGCAACGCGCGCGAGGTCCGACGCTGACCTGCGAGTGGCGGACAGGCGCGGA 1555
QY 406 -----Trp- 406
Db 1556 GCACCTGTGTGGCGAGATCGCTACTCGTCCCTTCGCGCGCGTGCAGCGCTGACCTGG 1615
QY 407 -----MetThrValGlnTyrAspAsnMetGlyArgValValLysGluL 422
Db 1616 GCAACGGCATCATCTCAGCGCGGAGTACGACGAGGACTACCAACTGCTCGGCGAG--- 1671
QY 422 euLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyG 442
Db 1672 --AAGTCGCGCGCTGCGAGAGC-----GACTACCAAGCAGATGCCAATGGCA 1717
QY 442 lnLeuGln-----ThrValSerIleAsnAspLysProLeuT 454
Db 1718 ATATCCAGAGACCGCCACAGCTCTGCGGACCCCTGGACTACGATCAGACCGCTGG 1777
QY 454 rpArg-----TyrSerTyrAspLeuAsnG 462
Db 1778 ACCGCTGACGAGGAGCGGCGTCCAGCGGCGGCGAGCTACGCTACGACGCGGTG 1837
QY 462 lyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspI 482
Db 1838 GCAACCGCACCAACGAGCGGACAAACCCCGCTCAGCGCGGCGACCGCGAGCGACT 1897
QY 482 leArg-----AspArgIleThrArgLeuGlyAspValGlnTyrLysMetA 497
Db 1898 ACCAGTACGCGCGCGACAGCAGCAGTACCGGCGCGCGCGCGCGCGCGCGCGCG 1957
QY 497 spGluAspGlyPheLeuArgGln---ArgGlyGlyAspIlePheGluTyrAsnSerAlaG 516
Db 1958 ACGCGCGCGCGCACTCACCAGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2017
QY 516 lyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspG 536
Db 2018 GCGCGCTG---CAAAGCGTCAGCTCGACGCGCGCGCGCGCGCGCGCGCGCGCG 2074
QY 536 lyLeuGlyArgArgVal 541
Db 2075 CCTTCGCGCGCGCGCATC 2091
```

RESULT 12

```
US-09-252-991A-7069
; Sequence 7069, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7069
; LENGTH: 2925
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7069

Alignment Scores:
Pred. No.: 4 17e-12 Length: 2925
Score: 182.50 Matches: 115
Percent Similarity: 36.06% Conservative: 88
Best Local Similarity: 20.43% Mismatches: 203
Query Match: 6.41% Indels: 157
DB: Gaps: 26
```

```
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-252-991A-7069 (1-2925)

QY 106 TyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArgLeuAsnGlyVal--- 124
Db 493 TATGACGAGCGCGCAAGCCCGCGCAGCATCACCGATGCC-----AACCGCGTCACC 543
QY 125 ---AsnValThrTyrSer---ProGlyGlyTyrIleAlaGlyIleGlnArgGlyIleMet 142
Db 544 AGCAGCTTCGCTTACACCGCGCGTGCAGCGGTGGCTGGCTTCCTCGTACGACCGCGCGCAGC 603
QY 143 SerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAlaAspGly 162
Db 604 ACCACCGCTTCGATCATGACGCGTGCAGCGCGCAGATCACCGCGCTCACCCGCGCGCGCGC 663
QY 163 LysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSerGlnArg 182
Db 664 AGTTGGCTGAGCTACGAATACGACGACGCGCGCGCTGTAGCGATCGCGCAACACCTC 723
QY 183 GluTyrIlePheGluPheAsp-----LysAsnAspArgLeu----- 194
Db 724 GCGCAACGCTCGAATACGACGCTGCACACCAAGGCAATCGCACCGCGCGCGCGCGCATCAAG 783
QY 195 -----SerSerValThrMetProAsnValAlaArgGlnThrLeu---GluThr 209
Db 784 GACGCCAGCGCGAGCTGTGCGCGCAGCAGCAATGGCGCTTACGACGAGCTTGGCGCGGTG 843
QY 210 IleArgSerValGly-----TyrTyrArgAsnIleTyrGlnPro 222
Db 844 CTCGTCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 903
QY 223 ProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThr 242
Db 904 CCGTTCGCGGAA-----ACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 942
QY 243 PheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLysLeuAla 262
Db 943 CAGCGCTTCGATGCGCTCGACCGCGTGTGCGCGCAGCGCATCCCTCGCGCGCGCGCG 1002
QY 263 GluThrLeuTyrAsp----- 268
Db 1003 CGACTCGCTACGACGCGCGCGCGCAACCTCAACGAGGTCAAGGACCGCGCGGTGTACCC 1062
QY 269 ThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGln 288
Db 1063 ACCCGC-----TACGAATACGCGCGCTCGGCAACCTGACCGCGCGCGCGCGCG 1110
QY 289 AsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGln 308
Db 1111 CCGGACGAGCGCGCACCCACCTTCGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCG 1170
QY 309 IlePheArgPheThrGluGluGlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsn 328
```



```
Db 1171 GACGTCGC-----|||:::|||||GGTGGGTCCAGGATCGCTACGACGCC 1209
Qy 329 SerPheArgValThrSerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeu 348
Db 1210 CTCATCGCTGCTCGAGCGCGCTCGCG--AGCGACCGAGCCTCGACGTACAG---1263
Qy 349 TyrArgTyrAspValSerGlyIleThrGluPheGlyIlePheGlyIleValIle---367
Db 1264 TACCGTACGACCTCACCGCGCGACGACACAGGCGATCGCGCGCTGGCGGCCATCGAT 1323
Qy 368 -----TyrTyrAsp-----|||:::|||||IleAsnGlnIle 374
Db 1324 GCGGTCGCGACAGCCTGTGTACCGCTACGACGAGCGGCGGCACTGTGCGAGCGTG 1383
Qy 375 IleThr-----|||:::|||||ThrAlaValMetThrHisThrIleHisPheAspAla 388
Db 1384 CGCAGCATCGGCTCGACGACGACGACCTGCTCGACCGCGGTGACCTACCGTACGACGCG 1443
Qy 389 TyrGlyArgMetLysGluValGlnTyrGluIlePheArgSerLeuMetTyr-----405
Db 1444 GCGAACCAACTGTGAGATCGGCTACCCCTCGCGCTCGCCATCGCTACCGCGCAAC 1503
Qy 405 -----|||:::|||||-----405
Db 1504 GCGGCGCGCAGGTCCGACGCTGACCTGGGCGAGTGGCGACAGGCGCGGACCCCTG 1563
Qy 406 -----|||:::|||||-----Trp-----406
Db 1564 GTCGGGCAGATCGCTACCTGCGCTCGCGCGCTGCTGCGCTGACCTGGGCGCAACGCG 1623
Qy 407 -----MetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysVal 424
Db 1624 ATCACTCTCAGCGCGAGTACGACGAGGACTACCACTCGCTCGCGAG-----AAGGTC 1677
Qy 425 GlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGln 444
Db 1678 GCGCCCTGCGAGC-----GACTACGACGACGATGCCAATGGCAATATCCAG 1725
Qy 445 -----ThrValSerIleAsnAspLysProLeuTrpArg---455
Db 1726 CAGCACCGCCACAGCCTCTCGGCGCACCTGGACTACGACGCGCTGGACCGCCTG 1785
Qy 456 -----TyrSerTyrAspLeuAsnGlyAsnLeu 464
Db 1786 ACCGAGGAACGGCGTCCAGGCGGCGCGAGCTATGCTTACGACGCGGTGCGCAACCGC 1845
Qy 465 HisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArg---483
Db 1846 ACCCAACGACGACACACCGCGCTCAGGCGGCGACCGCGAGCGAGCTACCACTAC 1905
Qy 484 -----AspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db 1906 GCGCCCGACACGACCGATTTGACCGCATCGCGCGCAACGCGTACGACGCGCGCGC 1965
Qy 500 GlyPheLeuArgGln---ArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeu 518
Db 1966 GGCAACTCACCAGACCGCGCGCGCGCGCAAACTGGCTTACGACCGCGCGCGCGCGCTG 2025
Qy 519 IleLysAlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGly 538
Db 2026 ---CAAGCGTCAGCTCGACGCGCAGCGGTGCGCGCATACGCTACCAACGCGCTCGC 2082
Qy 539 ArgArgVal 541
Db 2083 CAGCGCATC 2091
```

RESULT 13

US-09-328-352-2478
; Sequence 2478, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2478
; LENGTH: 4884
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2478

Alignment Scores:

Pred. No.: 1,148-11 Length: 4884
Score: 182.00 Matches: 137
Percent Similarity: 30.8% Conservative: 89
Best Local Similarity: 18.72% Mismatches: 198
Query Match: 6.40% Indels: 308
DB: 4 Gaps: 34

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-328-352-2478 (1-4884)

```
Qy 80 LeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLys 99
Db 1822 ATGAACCTAAATGGATGGTATTTTACCTACTTCAAAAGCCATTGAAGAATGGGCC---1878
Qy 100 PheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeu-----TrpSerPro 117
Db 1879 -----GACAATGCTAGTCTGTCTCAAGCTGAGTGGGATAAA 1917
Qy 118 SerSerArgLeu-----AsnGlyValAsnValThrTyrSer 129
Db 1918 AATATCCGAAACAGACAGTGTAGACGTTGAGGGAATAGCACTGAACATTACTACGAT 1977
Qy 130 ProGlyGlyTyrIleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGlu-----147
Db 1978 ATTGATGGCTAT-----ACTTACCGTATTGTTTATCTCTGATAATTTTGAAGAGTGT 2028
Qy 148 -----TyrAspGlnAlaGlyArgIleThrSerArgIlePheAlaAspGlyLysThrTrp 165
Db 2029 TTCTTTTCGAGATGATGCGAAATAATATACATCCATATTCTAAAGACGTTTCGAAACT 2088
Qy 166 SerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSerGln-----ArgGln 183
Db 2089 TCTTACACTTATGATGAGCGGTGATGTTTGTGACTCACTCACTCAAGATGATGTCGAC 2148
Qy 184 TyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetProAsnValAla 203
Db 2149 AGTTATTTTGAATATGATGATAAAAAAATCACTCACTGTTGATGTCGCCGACAGGGA 2208
Qy 204 Arg-----204
Db 2209 CGTTGGTTTAAACAATATGATGTTCTGGAAACCTGATCAAGAAATTTGATCCACTAAG 2268
Qy 205 ---GlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnProPro 223
Db 2269 CATGAACCTGCTTATGTTTACATGCGATGGCGGTAGTCCTTCCATACAGATGCTAAA 2328
Qy 224 GluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPhe 243
Db 2329 GGTGGAAGTAAGTCACTTAAATATGAT-----GATCAAGGGAATTTAATT---AGTTAT 2379
Qy 244 TyrLeuGlyThrGlyArgArgValIleTyrLysTyr-----GlyLysLeuSerLys 260
Db 2380 ACAGATTGCTCAGGTAAGAGACCAAGTGGCAATATGATGAGCGGTGCTGTGTAATAGC 2439
Qy 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrThrThrThrAla---279
Db 2440 ATTGAATAATGCATTA-----AACCAAAAGTTGAGTATTTTATCTGAACCTACCTTA 2493
Qy 280 -----GlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThr 295
Db 2494 GAAATCGGGAACCAATTATTAAGGGTTTACCTCTGAATGCTTTTGTTCACACTTGAAGAAG 2553
```

Qy 296 IleArgTyr----- 298
Db 2554 ATTAACATGCTGATGGCACAGAACACCTTCATTCATGATGCTGAAGGGCGTTTACTG 2613
Qy 299 ArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetVal 318
Db 2614 GCGCATGTAGATCTTAAACAGAACATCACTCGCTATGAATATGATGAGCGAGGTTTAAAT 2673
Qy 319 AsnAlaArgPheAsp-----TyrAsnTyrAspAsnSerPheArg 331
Db 2674 CTGCTCTGTACAGATGCTCTTAAACATAAATATAAATATAAATGGGATCGAATGGGACGT 2733
Qy 332 ValThrSerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyr 351
Db 2734 CTCACCGCACTC-----ATTAAAGAAATGGTGCCAGTTACCAATTCCTC---TAT 2781
Qy 352 AspAspValSerGlyLysThrGluGlnPhe-----GlyLysPheGlyValIleTyr 368
Db 2782 GATGTGCGAAGCGGACTCGTTTAAAGAAATTCACCTTTGATGGTAAGGAAACGGTTTACCAT 2841
Qy 369 TyrAsp-----IleAsnGlnIleThrThrAlaValMetThrHisThrLysHisPhe 386
Db 2842 TACGATGAAAGTCGGGCACTTCGCAACCACTGATGAAGTTGCA----- 2886
Qy 387 AspAlaTyrGlyArg--MetLysGlu-----ValGlnTyrGlu 398
Db 2887 TCTGCTATGCAAGACCTCAAGACCGAGCTGCACGAAGACCGAATTCACAAATTT 2946
Qy 399 IlePheArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValVal 418
Db 2947 ATTTT-----GACAGTATGGGCGTTTAGAA 2973
Qy 419 LysLysGluLeuLysValGlyProTyrAla-----AsnThrThrArg 432
Db 2974 CAGCGTACAGCGGTTATGGTCATTTAGTGTGTTAGAACTGAAGAAAAACAGACTGAAGAA 3033
Qy 433 TyrSerTyrGlu----- 436
Db 3034 TTTCCTATGATTATATGGGCGAATCATTCAGCTAAATATGCTCAAGTAATTTACAA 3093
Qy 437 -----TyrAspAlaAspGlyGlnLeu-----GlnThrValSerIleAsn 449
Db 3094 TGGTTTATGATGCGGCTGGCAATTTAGTACAGAAACACCAAGCAATTTATAAATCAAT 3153
Qy 450 AspLysProLeuTyrArgTyrSerTyrAsp----- 459
Db 3154 AAACACGGGTGTGGAAACCAATATGACGAAATTAATGATCGAATCAAAACAACTCGT 3213
Qy 459 ----- 459
Db 3214 CCAGACGGCAGGTTATTGACTGTTGACCTATGCTAGCGTCAATGACAAAGTTTAATA 3273
Qy 460 LeuAsnGly-----AsnLeuHis----- 465
Db 3274 GTGAATGCTCAAGACTTTGTTAGCTTTGAGCGTGATGATTTACACCGTGAATCGCAGT 3333
Qy 465 ----- 465
Db 3334 CACTATGCCACGAGTGAGTCAGAACAGCAATATGACTTAGCAGGTCGTTTGAAGAAGC 3393
Qy 466 -----LeuLeuSerPro-----GlyAsn 471
Db 3394 CAATATGATGTTAAGTGAACATGAATAAATGGCTATCAAAACCAATATAAGCGCCACAAAT 3453
Qy 472 SerAlaArgLeuThrPro-----LeuArgTyrAsp----- 481
Db 3454 CATTAGACGACAGCTCAACTCGTACAGGCTTTTATCAATATGACAAACAGAGGTGAG 3513
Qy 482 -----IleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGlu 498
Db 3514 CTAACGGCAATTCAGAT-----ACACGGGTGGCAATATTGCTTATTAATACGATCCT 3567

Qy 499 AspGlyPheLeuArgGlnArg-----GlyGlyAspIlePheGluTyrAsnSer 514
Db 3568 GTAGGGCGTTTACTTGAAGCGCAGTAAGTTAGTAAAGAAACATTTAGTTTGACCT 3627
Qy 515 AlaGlyLeuLeuIleLysAlaTyr----- 522
Db 3628 GCGAGTAACATTTCTAGATTTCATACCATAGTCAGAAAGTTCAAAGCCATTCACAAAGCTA 3687
Qy 522 ----- 522
Db 3688 GACGAAACGAGTTATGTTTAAATCGTTTGGTTTAAATAGTAGTTAAAGAAATATTAGAC 3747
Qy 522 ----- 522
Db 3748 CAACAATACCAATATGCTTATGGCAACTCATACGACAAAAACGAGCAAGGTGAT 3807
Qy 523 -----AsnArgAlaGlySerTyrSer 529
Db 3808 TTAACCTTGAGTGGGATGCTATGGCGCATGGTAAAGAGCGGTAATAGTCAATACACG 3867
Qy 530 ValArgTyrArgTyrAspGlyLeuGlyArgVal 541
Db 3868 GCAGATATCGTTATGATGCTTTAGGCGCGCGTATT 3903
RESULT 14
US-09-252-991A-11787
; Sequence 11787, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11787
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11787
Alignment Scores:
Pred. No.: 8,26e-12 Length: 1857
Score: 177.50 Matches: 113
Percent Similarity: 35.44% Conservative: 89
Best Local Similarity: 19.82% Mismatches: 196
Query Match: 6.24% Indels: 172
DB: 4 Gaps: 24
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-252-991A-11787 (1-1857)
Qy 54 TrpArgGlnArgLysGluGlnAlaArgGlyGlnValThrValPheGlyArgArgLeuArg 73
Db 31 TGGCGCGGCGATGGCTGAG-----CTACGA 57
Qy 74 ValHisAsnArgAsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIle 93
Db 58 ATACGACGACGACGCGCGCTGGTGGCGATCGGCAACAACTCGCGCAAGCTCTCGAATA 117
Qy 94 TyrAspAspHisArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSer 113
Db 118 CGAGCTGCACCA-----GGCAATCGACCGCCGCGG 153
Qy 114 -LeuTyrSerProSerSerArgLeuAsnGlyValAsnValThrTyrSerProGlyGly 133
Db 154 CATCAAGGACGCGACGCGCGCTG----- 178
Qy 133 rIleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArg 153

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: August 14, 2004, 01:13:26 ; Search time 6344.74 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_1760_2300
Perfect score: 2845
Sequence: 1 YYIGADGSLRLLLANGNEVA.....YNRAGSWSVRYDGLGRV 541

Scoring table:
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_13135 @runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_invm:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|----------|---------------------|
| 1 | 2834 | 99.6 | 6246 | 29 | AY413475 | AY413475 Homo sapi |
| 2 | 2829 | 98.4 | 5970 | 29 | AY413476 | AY413476 Pan trogl |
| 3 | 2233 | 72.5 | 5090 | 29 | AY405420 | AY405420 Homo sapi |
| 4 | 2228 | 78.3 | 5069 | 29 | AY405421 | AY405421 Pan trogl |
| 5 | 2228 | 78.3 | 5087 | 29 | AY405422 | AY405422 Mus muscu |
| 6 | 2017.5 | 70.9 | 3038 | 11 | AK037897 | AK037897 Mus muscu |
| 7 | 1177 | 41.4 | 894 | 13 | BU151768 | BU151768 AGENCOURT |
| 8 | 1174 | 41.3 | 834 | 13 | BU610605 | BU610605 UI-M-FCO- |
| 9 | 1168.5 | 41.1 | 951 | 13 | BU839812 | BU839812 AGENCOURT |
| 10 | 1142 | 40.1 | 694 | 14 | CF723353 | CF723353 UI-M-GVO- |
| 11 | 1112.5 | 39.1 | 758 | 14 | CB723840 | CB723840 UI-M-FYO- |
| 12 | 1103 | 38.8 | 757 | 10 | BE741784 | BE741784 601595546 |
| 13 | 1099 | 38.6 | 868 | 29 | CNS03W3H | AL263222 Tetraodon |
| 14 | 1097 | 38.6 | 856 | 14 | CF534506 | CF534506 UI-M-G10- |
| 15 | 1068 | 37.5 | 1031 | 29 | CNS02PHB | AL207992 Tetraodon |
| 16 | 1011 | 35.5 | 675 | 14 | CF540583 | CF540583 UI-M-GVO- |
| 17 | 1005 | 35.3 | 805 | 13 | BU215426 | BU215426 603756042 |
| 18 | 1000.5 | 35.2 | 868 | 14 | CA454304 | CA454304 AGENCOURT |
| 19 | 999 | 35.1 | 812 | 9 | AU133387 | AU133387 AU133387 |
| 20 | 992.5 | 34.9 | 728 | 14 | CF735545 | CF735545 UI-M-HB0- |
| 21 | 986 | 34.7 | 600 | 14 | CA527167 | CA527167 8032-72 M |
| 22 | 983 | 34.6 | 690 | 28 | AZ329056 | AZ329056 IM0053E09 |
| 23 | 979 | 34.4 | 811 | 9 | AU119933 | AU119933 AU119933 |
| 24 | 961 | 33.8 | 655 | 14 | CD351410 | CD351410 UI-M-G10- |
| 25 | 954 | 33.5 | 1048 | 29 | CNS03Z5U | AL267195 Tetraodon |
| 26 | 938 | 33.0 | 819 | 13 | BQ443891 | BQ443891 UI-M-EW0- |
| 27 | 920 | 32.3 | 944 | 29 | CNS02TBX | AL212874 Tetraodon |
| 28 | 896 | 31.5 | 775 | 13 | BU611301 | BU611301 UI-M-F10- |
| 29 | 888 | 31.2 | 751 | 14 | CF535325 | CF535325 UI-M-GHO- |
| 30 | 885 | 31.1 | 764 | 13 | BU328309 | BU328309 603495004 |
| 31 | 876 | 30.8 | 592 | 10 | BF349559 | BF349559 MR0-HT015 |
| 32 | 874 | 30.7 | 668 | 13 | BU214946 | BU214946 603108244 |
| 33 | 860 | 30.2 | 763 | 29 | CNS03WBL | AL263494 Tetraodon |
| 34 | 844 | 29.7 | 589 | 14 | CB434224 | CB434224 610820 MA |
| 35 | 837 | 29.4 | 882 | 13 | BU914563 | BU914563 AGENCOURT |
| 36 | 831.5 | 29.2 | 900 | 13 | BU267121 | BU267121 603506284 |
| 37 | 826 | 29.0 | 675 | 10 | BE374675 | BE374675 601224176 |
| 38 | 817 | 28.7 | 1103 | 13 | EX457474 | EX457474 BX457474 |
| 39 | 808 | 28.4 | 633 | 14 | CF728295 | CF728295 UI-M-HB0- |
| 40 | 799 | 28.1 | 577 | 29 | DR7F15T | AL740238 Danilo rer |
| 41 | 778 | 27.3 | 748 | 14 | CF723503 | CF723503 UI-M-GVO- |
| 42 | 765 | 26.9 | 444 | 14 | CB546578 | CB546578 AMGNNUC:N |
| 43 | 764 | 26.9 | 843 | 9 | AU124680 | AU124680 AU124680 |
| 44 | 763.5 | 26.8 | 592 | 10 | BE161779 | BE161779 MR3-HT044 |
| 45 | 763 | 26.8 | 974 | 13 | BQ425244 | BQ425244 AGENCOURT |

ALIGNMENTS

RESULT 1
AY413475
LOCUS
DEFINITION Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
6246 bp DNA linear GSS 12-DEC-2003
ACCSSION AY413475
AY413475 genomic survey sequence.
VERSION AY413475.1 GI:39769437
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 6246)


```

Db      4828 CGCGCGGTG 4836
|||||
RESULT 2
LOCUS   AY413476
DEFINITION Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY413476
VERSION   AY413476.1 GI:39769438
KEYWORDS  GSS.
SOURCE    Pan troglodytes (chimpanzee)
ORGANISM  Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 5970)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 5970)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (18-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES             source
    source            1..5970
                        /organism="Pan troglodytes"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9598"
                        <1..>5970
                        /locus_tag="HCM4903"
ORIGIN
Alignment Scores:
Pred. No.:          2,27e-310      Length:          5970
Score:              2629.00         Matches:          505
Percent Similarity: 93.19%          Conservative:     1
Best Local Similarity: 93.00%        Mismatches:      35
Query Match:        92.41%          Indels:          2
DB:                 29              Gaps:            1
US-10-029-020-14_COPY_1760_2300 (1-541) x AY413476 (1-5970)
QY      1   TyrTyrileGlyAlaaspGlySerLeuArgLeuLeuLeuAlaasnGlyMetGluValala 20
Db      2932 TACTACATCGGGCCCGATGGCTCTCTGCGCTGCTGCTGCGCCACCGCATGGAGGTGGT 2991
QY      21   LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db      2992 CTCGAGACTGAGCCCACTTGTGCTGGTGGCCACCGTCAACCCCGTGGGCAAGAGGAAT 3051
QY      41   ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluThrArgGlnArgLysGluGln 60
Db      3052 GTCACGCTGCCATCGACAAACGGCTCAACCTGGTGGAGTGGCCAGCGCAAGAGCAG 3111
QY      61   AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db      3112 GCTCGGGGCGAGGTCACTGTCTTTGGCGCGGCTCGCGGTTCACAAACCGAAATCTCCTA 3171
QY      81   SerLeuAspPheAspArgValThrArgThrGluLysValleThrAspAspHisArgLysPhe 100
Db      3172 TCTCTGGAATTTGATCGCGTACACCGACAGAGAGATCTATGATGACCCGCAAGTTC 3231
101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyLysGlnArgGly 140
3292 CTGAATGGTGTCACTGACATACCTCCCTGGAGGTATCACTTGTGGCATCCAGAGGGC 3351
141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
3352 ATCATGCTCTGAAGAATGAATACGACAGCGGGCGCGCATCATCCAGGATCTTCGCT 3411
161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSer-----MetValLeuLeuLeu 178
3412 GATGGGAAGACATGGAGCTACACATACTTAGAAGAGCGAGTNNNNNNNNNNNNNNNNNN 3471
179 HisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThr 198
3472 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3531
199 MetProAsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsn 218
3532 ATGCCAAACGTGGCGCGCAGACNNNAGAGACCATCCGCTCAGTGGGCTTACTACAGAAC 3591
219 IleTyrGlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHis 238
3592 ATCTATCAGCCCGCTTGGAGCAATGCCCTCAGTCATACAGGACTTCCTCAGGATGGGCAC 3651
239 LeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeu 258
3652 CTCCTTCACACTTCTACCTGGGCACTGGCGCGAGGTGATATACAGTATGGCAAACTG 3711
259 SerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThr 278
3712 TCAAAGCTGGCAGAGACGCTCTATGACACACCAAGGTGCTTTCCTATGACGAGACG 3771
279 AlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyr 298
3772 CGAGGCATGCTGAAGACCATCACTACAGAAATGAGGCTTCACTGCACCATCCGCTAC 3831
299 ArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetVal 318
3832 CGTCAGATTGGGCCCCCTGATTGACCGACATCTTCGCTTCCCGGAGGAGGATGGT 3891
319 AsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaVal 338
3892 AACGCCCGTTTTCGACTACAACTATGACACAGCTTCGGGTGACCACTGCAGGCTGTG 3951
339 IleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThr 358
3952 ATCAACGAGACCCCACTGCCCATTTGATGCTATCGCTATGATGATGTGTGAGCAAGACA 4011
359 GluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleThrThrAla 378
4012 GACCACTTTGGGAAGTTTGGTGTCACTTATGATTAACCAAGATCACTACCAAGCT 4071
379 ValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGlu 398
4072 GTCATGACCCACACCAAGCATTTTGTATGATATGCGAGATGAGGAAGTGCAGTATGAG 4131
399 IlePheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVal 418
4132 ATCTTCGCTCGCTCATGTACTGAGTACTGTCCAGTATGATTAACCAAGTGCAGTATGAG 4191
419 LysLysGluLeuLysValGlyProTyrAlaAsnThrArgTyrSerTyrGluTyrAsp 438
4192 AAGAGGAGCTGAAGTAGGACCCCTACGCCAATACCACTCGCTACTCTATGATGATGAT 4251
439 AlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyr 458
4252 GCTGAGCGCCAGCTGCAGACAGCTTCCATCAATGACCAAGCCACTCTGGCGCTACAGTAC 4311
459 AspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeu 478

```

Mon Aug 16 09:01:11 2004

| | | | |
|---|------------|--|------|
| Db | 4312 | GACCTCATGGGAACCTGCACTTACTGAGCCCTGGGAAACAGTGCAGCGCTCACACCACCTA | 4371 |
| Qy | 479 | ArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGlu | 498 |
| Db | 4372 | CGGTATGACATCCGGCAGCGATCACTCGGCTGGGTGAGCTGCAATCAAGATGATGAG | 4431 |
| Qy | 499 | AspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeu | 518 |
| Db | 4432 | GATGCTTNTTGGAGCAGCGGGCGGTGATATCTTTGAGTACAACCTCANCTGNNNTANNC | 4491 |
| Qy | 519 | IleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGly | 538 |
| Db | 4492 | ATCAAGGNCNACACCGGGCTGGCAGCTGGAGTGTGAGTACCGTACGATGNNNTGGGG | 4551 |
| Qy | 539 | ArgArgVal 541 | |
| Db | 4552 | CGCGCGGTG 4560 | |
| RESULT 3 | | | |
| LOCUS | AY405420 | 5094 bp DNA linear GSS 12-DEC-2003 | |
| DEFINITION | | Homo sapiens HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence. | |
| ACCESSION | AY405420 | | |
| VERSION | AY405420.1 | GI:39761394 | |
| KEYWORDS | | GSS. | |
| SOURCE | | Homo sapiens (human) | |
| ORGANISM | | Homo sapiens | |
| REFERENCE | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| AUTHORS | | 1 (bases 1 to 5094) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | |
| TITLE | | Inferred nonneutral evolution from human-chimp-mouse orthologous gene trios | |
| JOURNAL | | Science 302 (5652), 1960-1963 (2003) | |
| PUBMED | | 14671302 | |
| REFERENCE | | 2 (bases 1 to 5094) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | |
| AUTHORS | | Direct Submission | |
| TITLE | | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | |
| JOURNAL | | | |
| COMMENT | | This sequence was made by sequencing genomic exons and ordering them based on alignment. | |
| FEATURES | | Location/Qualifiers | |
| source | | 1..5094 | |
| | | /organism="Homo sapiens" | |
| | | /mol_type="genomic DNA" | |
| | | /db_xref="taxon:9606" | |
| gene | | <1..5094 | |
| | | /locus_tag="HCM2218" | |
| ORIGIN | | | |
| Alignment Scores: | | 8,47e-262 Length: 5094 | |
| Pred. No.: | | 2233.00 Matches: 404 | |
| Score: | | Conservative: 77 | |
| Percent Similarity: | | 88.91% Mismatches: 60 | |
| Best local Similarity: | | 74.68% Indels: 0 | |
| Query Match: | | 78.49% Gaps: 0 | |
| DB: | | 29 | |
| US-10-029-020-14_COPY_1760_2300 (1-541) x AY405420 (1-5094) | | | |
| Qy | 1 | TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20 | |
| Db | 2381 | TACCAGATTGGTTATCAGCGCTCCCTCAGATTATATCTACGCCAGTGGCTGGACACAC 2440 | |
| Qy | 21 | LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40 | |
| Db | 2441 | TACCAAAACAGAGCGCACGTTCTGGCTGGCACCGCTAATCCGACGGTGGCCAAAGAAAC 2500 | |
| Qy | 41 | ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60 | |
| Db | 2501 | ATGACTTTTGGCTGGCGAAGCGGTCAAACCTTGGTGGATGGAGATTCCGAAAGAGCAA 2560 | |
| Qy | 61 | AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80 | |
| Db | 2561 | GCCCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGTTAATGGGCAAGAACCTCCTT 2620 | |
| Qy | 81 | SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAsnHisArgLysPhe 100 | |
| Db | 2621 | TCAGTTGACTTTGATCGAACACAAAGACAGAAAAGATCTATGACGACCCGCCGTAATTT 2680 | |
| Qy | 101 | ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerArg 120 | |
| Db | 2681 | CTACTGAGGATCGCTTACGACAGTCTGGGACCCGACCTCTCTGGTGCCTCAAGCAGCAAG 2740 | |
| Qy | 121 | LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140 | |
| Db | 2741 | CTGATGGCGCTCAATGTCACTTATTCATCCACAGGTCAAATTTGCCACGATCCACGCGGC 2800 | |
| Qy | 141 | IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160 | |
| Db | 2801 | ACCACCTAGCGAGAAAGTAGATTATGACGGACAGGGGAGGATCGTGTCTCGGTCTTTGCT 2860 | |
| Qy | 161 | AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180 | |
| Db | 2861 | GATGGTAAACATGAGGAGTTACACATATTTAGAAAAGTCCATGGTCTCTCTGCTTCATAGC 2920 | |
| Qy | 181 | GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200 | |
| Db | 2921 | CAGCGCAGTACATCTTCAATACGATATGTGGACCGCTGTCTCCATCCATCCATGCCCC 2980 | |
| Qy | 201 | AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220 | |
| Db | 2981 | AGTGTGGCTCGCCACACCATCGAGCATCCGATCCATTCGCTACCTACCCCAACATATAC 3040 | |
| Qy | 221 | GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240 | |
| Db | 3041 | AACCCCGGAAAGAACCGCTCACTCATCCGAGTACACGAGGAGGAGGCTGCTCTTA 3100 | |
| Qy | 241 | HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260 | |
| Db | 3101 | CAAAACAGCTTCTTGGGTACAAGTCGAGGGTCTTATTCAATACACAGAGCAGACTAGG 3160 | |
| Qy | 261 | LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280 | |
| Db | 3161 | CTCTCAGAAAATTTTATATGATAGCACAGAGTCAGTTTTTACCTATGATGAACAGCAGGA 3220 | |
| Qy | 281 | MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300 | |
| Db | 3221 | GTCCTAAAGACAGTAACCTCCAGATGATGTTTATTGTCACCAITAGATACAGGCAA 3280 | |
| Qy | 301 | IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320 | |
| Db | 3281 | ATTGGTCCCTGATTGACAGGAGATTTTCGCTTTAGTGAAGATGGATGGTAAATGCA 3340 | |
| Qy | 321 | ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340 | |
| Db | 3341 | AGATTTGACTATAGTATGACACAGCTTTCGATGACACGATGACGAGGTGTGATCAAT 3400 | |
| Qy | 341 | GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360 | |
| Db | 3401 | GAACGCCACTGCCTATTGATCTGTATCAGTTATGATGACATTTCTGGCAAAAGTTGAGCAG 3460 | |
| Qy | 361 | PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380 | |
| Db | 3461 | TTTGGAAAGTTTGGAGTTTAT 3520 | |
| Qy | 381 | ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400 | |

Db 3034 AATCCCCAGAAAGCAATGCTCTATCATCCGACTACACGAGGAGGGCTGCTCTG 3093
Qy 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 3094 CAACAGCTTTCTCTGGGAACGAGTCGGAGGGCTTTATTCAAGTATAGAAGGAGCAGCAGG 3153
Qy 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 3154 CTATCAGAAATTTTATACGACGACACAGAGTCAGTTTACCTACGACGAAACACGGGGA 3213
Qy 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 3214 GTCCTGAAACAGTAACCTTCAGAGTGATGGTTTATTTCACCATAGATACAGGCAA 3273
Qy 301 IleGlyProLeuIleAsnArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 3274 ATTGTCCCTGATTCAGACAGAGATTTTCGGCTTCAGCGAGGATGGAATGTAATGCG 3333
Qy 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 3334 AGATTGACTATAGTACGACAAACAGCTTTCAGTGACACGATGACGGGTGTCAATCAAT 3393
Qy 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
Db 3394 GAAACACCACTGCCCAATGATCTATACCAAGTTTGATGACATCTCTGGCAAGTCGACGAG 3453
Qy 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 3454 TTTGGAAATTCGGAGTGATATACAGACATCAACCAAAATCATTTCCACGCCGCTGATG 3513
Qy 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 3514 ACTTATACAAAGCACTTTGATGCTCATGGCGCATCAAGGAGATCCCAATATGAGATATTT 3573
Qy 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 3574 AGGTCACCTCATGTACTGATTCACAAATCAATATGATGATGCGCGGGTAACCAAGAGA 3633
Qy 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 3634 GAGATTAAATTTGGGCCCTTTTGGCAACCACTACCAAAATACCGGTACGAGTACGATCGAT 3693
Qy 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
Db 3694 GGACAGCTCCAAACAGTTTACCTTAAACGAAAGATCATGTGGCGGTACAACTACGACCTA 3753
Qy 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 3754 AATGGAAACCTCCACTTGCTCAACCCAGCAGCAGCGCCGCTGACCCCTCTGCGCTAT 3813
Qy 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 3814 GACCTGGCGACAGATCACCCGCTGGCGGATGTTCAGTACCGCTGGATGAAGATGGT 3873
Qy 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLysLys 520
Db 3874 TTCCTCGCTCAGAGGGGCACTGAAATTTTGAATACAGCTCCAAAGGGCTTCGACTCGA 3933
Qy 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 3934 GTCTACAGTAAGAGCGATGGCTGGCAGTATCTATCGGTACGACGCGCTGGGAAGACGT 3993
Qy 541 Val 541
Db 3994 GTT 3996

RESULT 6

AK037897

LOCUS

DEFINITION Mus musculus 16 days neonate thymus cDNA, RIKEN full-length enriched library, clone:AI30606124 product:odd Oz/ten-m homolog 1 (drosophila), full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK037897
AK037897.1 GI:26332325
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3038)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tanaka, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

REFERENCE
AUTHORS
TITLE
JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/

COMMENT

| | | | |
|----|-----|--|-----|
| QY | 18 | GluValAlaLeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGly | 37 |
| Db | 3 | GAGATCAACCTCAGCTCGAGGCCACATCTCTGGCAGGGGCTGTCAACCCCACTCTAGGC | 62 |
| QY | 38 | LysArgAsnValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTtpArgGlnArg | 57 |
| Db | 63 | AAATGCACCATATCGTTGCTCGAGGACACAAATGCAAACCTCATTTAGTGGAGACAGAG | 122 |
| QY | 58 | LysGluGlnAlaArgGlyGlnValThrPheGlyArgArgLeuArgValHisAsnArg | 77 |
| Db | 123 | ARGGACCAAAACAAGGCAACGTTTCAGCTTTTGAGAGAAGACTGAGGGCCCAATAACAGA | 182 |
| QY | 78 | AsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHis | 97 |
| Db | 183 | AACTCTATTTCCTGATTTTCATCATGACCCGACGGGTAGATCTATGATGACCAC | 242 |
| QY | 98 | ArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerPro | 117 |
| Db | 243 | CGAAATTCCTCTTCGAATTCCTTATGACCAGACCGGGCGACCTATTCTGTGCTCTCT | 302 |
| QY | 118 | SerSerArgLeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIle | 137 |

```

QY 498 GluAspGlyPheLeuArgGlnArgGlyGlyAspPheGluTyrAsnSerAlaGlyLeu 517
Db 1440 GAAGAGCGTTTCTGAGACAGAGGGGAATGATATATTGAGTAAATTCCTAATGGTCTG 1499

QY 518 LeuIleLysAlaTyrAsnArgAlaGlySerTyrPheValArgTyrArgTyrAspGlyLeu 537
Db 1500 CTGCAGAAAGCCTCAATAAAGTTTCTGCTGGACTGTACAGTATTACTATGATGGCTT 1559

QY 538 GlyArgArgVal 541
Db 1560 GGGGACGCGTT 1571

RESULT 7
BUI51768
LOCUS BUI51768
DEFINITION AGENCOURT_8750792 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6335962
5', mRNA sequence.
ACCESSION BUI51768
VERSION BUI51768.1 GI:22665300
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 894)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
CDNA Library Preparation: ResGen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIA013799 row: d column: 11
High quality sequence stop: 677.
FEATURES
source
1. .894
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6335962"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_130"
/note="Organ: oocytes; Vector: pCMV-SPORT6.1; Site:1:
EcoRV; Site:2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.95 kb. Constructed by
ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 2,87e-133 Length: 894
Score: 1177.00 Matches: 208
Percent Similarity: 93.51% Conservative: 37
Best Local Similarity: 79.39% Mismatches: 17
Query Match: 41.37% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x BUI51768 (1-894)

QY 280 GlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArg 299
Db 3 GGAGTCCTGAAACAGTAAACCTTCAGAGGTGATGTTTATTGTCACCATAGATACAGG 62

QY 300 GlnIleGlyProLeuLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsn 319
Db 63 CAAATTGGTCCCTGATTGACAGACAGATTTCCGCTTCAGCGAGGATGGAATGGTAAAT 122

```

```

QY 320 AlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIle 339
Db 123 GCGAGATTGATATAGCTAGCACAACAGCTTCGAGTGACAGCATGCAGGGTGTCAATC 182

QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu 359
Db 193 AATGAACACCACTGCCCATTTGATCTATACCAAGTTTGATGACATCTCTGGCAAGTCGAG 242

QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379
Db 243 CAGTTTGGAAATTCGGAGTGTATATCTACGACATCAACCAATCATTTTCCAGGCCGCTG 302

QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db 303 ATGACTTATACAAACACTTTTGATGCTCATGGCGCATCAAGGAGATCCAATATGAGATA 362

QY 400 PheArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValLys 419
Db 363 TTTAGTCACTCATGTACTGATTTACAAATCAATGATGATATATGGCCGGGTACCAAG 422

QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db 423 AGAGAGATTAAATTTGGGCTTTTGGCCACACTACCAATAACGGTACGATACAGCTC 482

QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAsp 459
Db 483 GATGGACAGCTCCAAACAGATTACCTAAACGAAAGATCATGTGGCGGTACAACTACGAC 542

QY 460 LeuAsnGlyAsnLeuHisLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db 543 CTAATGGAAACCTCCACTTGTCTCAACCCAGCAGCAGCGCCGCTGACCCCTCTGGCC 602

QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db 603 TATGACCTGGCGGACAGAAATCACCGCTGGCGCATGTTTCAGTACCGCGCTGATGAAGAT 662

QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
Db 663 GGTTCCTCGCTCAGAGGGGCACTGAAATTTTGAATACAGCTCCAAAGGGCTTCTGACT 722

QY 520 LysAlaTyrAsnArgAlaGlySerTyrPheValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 723 CGAGTCTACAGTANAGCGAGTGGCTGACAGTGTATCTATCGTACGACGCGCTGGGAAGA 782

QY 540 ArgVal 541
Db 783 CGTGT 788

RESULT 8
BUI610605
LOCUS BUI610605
DEFINITION UI-M-FCO-caq-f-07-0-UI.r1 NIH BMAP_FCO Mus musculus cDNA clone
UI-M-FCO-caq-f-07-0-UI 5', mRNA sequence.
ACCESSION BUI610605.1 GI:23276820
VERSION BUI610605.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 834)
NIH-MGC http://mgi.nci.nih.gov/
REFERENCE National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS Unpublished (1999)
TITLE Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
This clone was contributed by the Brain Molecular Anatomy Project

```

| | | | | | |
|--------------------|--|---|--|--|--|
| (BMAP) | | Location/Qualifiers | | Db | |
| Seq primer: pYX-5. | | 1. .834 | | 482 | |
| FEATURES | | /organism="Mus musculus" | | CAGCTCCAAACAGTTTACCTAAACAAAGATCATGTGGCGGTACACTACGACCTAAAT | |
| source | | /mol_type="mRNA" | | 462 | |
| | | /strain="C57BL/6" | | GlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAsp | |
| | | /db_xref="taxon:10090" | | 542 | |
| | | /clone="UI-M-FCO-caq-f-07-0-UI" | | GGAAACCTCCACTTGTCTAACCCAGCAGCAGCCGCCCTGACCCCTCTCGCTATGAC | |
| | | /tissue_type="whole brain" | | 482 | |
| | | /dev_stage="embryo 12.5 dpc" | | lIeArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGlyPhe | |
| | | /lab_host="DH10B (T1 phage resistant)" | | 602 | |
| | | /clone_lib="NIH BMAP_FCO" | | CTGCCGACAGATCACCCGGCTGGCGGATGTTTCAGTACCGCTGGATGGATGGTTTC | |
| | | /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAAGAGACC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator." | | 502 | |
| | | | | LeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleYsala | |
| | | | | 662 | |
| | | | | CTGCGTCAGAGCGGCACCTAAATTTTGAATACAGCTCCANAGGCTTCTGACTCGAGTC | |
| | | | | 522 | |
| | | | | TyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyValArgVal | |
| | | | | 722 | |
| | | | | TACAGTAAAGGAGTGGCTGGACAGTATCTATCGTACAGCGGCTGGGAAGAGTGTT | |
| | | | | RESULT 9 | |
| | | | | BU839812 | |
| | | | | LOCUS | |
| | | | | DEFINITION | |
| | | | | 5', mRNA sequence. | |
| | | | | ACCESSION | |
| | | | | VERSION | |
| | | | | KEYWORDS | |
| | | | | SOURCE | |
| | | | | ORGANISM | |
| | | | | REFERENCE | |
| | | | | AUTHORS | |
| | | | | TITLE | |
| | | | | JOURNAL | |
| | | | | COMMENT | |
| | | | | FEATURES | |
| | | | | source | |
| | | | | 1. .951 | |
| | | | | /organism="Mus musculus" | |
| | | | | /mol_type="mRNA" | |
| | | | | /db_xref="taxon:10090" | |
| | | | | /clone="IMAGE:6329777" | |
| | | | | /lab_host="DH10B (phage-resistant)" | |
| | | | | /clone_lib="NIH_MGC_130" | |
| | | | | /notes="Organ: otocysts; Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 Kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library." | |
| | | | | ORIGIN | |
| | | | | Alignment Scores: | |
| | | | | Pred. No.: | |
| | | | | Score: | |
| | | | | Percent Similarity: | |
| | | | | Best Local Similarity: | |
| | | | | Query Match: | |
| | | | | DB: | |
| | | | | US-10-029-020-14_COPY_1760_2300 (1-541) x BU839812 (1-951) | |
| | | | | 143 | |
| | | | | SerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAlaAspGly | |
| | | | | 162 | |

```
Db 23 AGCGAAAGTGTGACTATACAGCCAGGAGGATCTATCTCGGGTCTTTGCCGATGG 82
QY 163 LysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSerGlnArg 182
Db 83 AAAACATGGAGTTACAGTACTTGGAAAGTCTCATGGTCTTCTGCTCATAGCCAGCG 142
QY 183 GlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetProAsnVal 202
Db 143 CAGTATCATCTTCAATACGACATGTGGGACCGCTCTCGGCCATCACCATGCCAGTGTG 202
QY 203 AlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnPro 222
Db 203 GTCGCCACACCATGACGACCATCGGTCCTCATGGCTACTACCGAACATCTACATCC 262
QY 223 ProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuHisThr 242
Db 263 CCAGAAAGCAATGCTTATCATCATCCGACTACAGGAGGAGGGTGTCTTGTCAAAACA 322
QY 243 PheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLysLeuAla 262
Db 323 GCTTTCTCGGACAGTCCGAGGTCTTATTCAAGTATAGAGGACAGCAGGCTATCA 382
QY 263 GluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeu 282
Db 383 GAAATTTATACGACAGACAGATGCTAGTTTACCTACGACGAAACAGCGGAGTCTGT 442
QY 283 LysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGly 302
Db 443 AAAACAGTAACCTTCAGAGTGATGTTTATTGCAACATTAGATACAGGCAATTGGT 502
QY 303 ProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAlaArgPhe 322
Db 503 CCCCTGATGTACAGACATTTTCGCTTCAGCGAGGATGAATGTTAAATCCGAGATT 562
QY 323 AspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsnGluThr 342
Db 563 GACTATGCTACGACACAGCTTTTCGAGTGACGACATGACGAGGTGTTCATCAATGAA 622
QY 343 ProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGlnPheGly 362
Db 623 CCACGCCCCATTGATCTATACCATTTGATGACATCTCTGGCAAGTCGACAGTTTGG 682
QY 363 LysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHi 382
Db 683 AAATTCGGAGTATATACATCAACCAATATCTTCCACGGGCGTGTGATCTTA 742
QY 382 sThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePheArg 402
Db 743 TACAAAGCATTTGATGCTCATGGCGCATCAAGGAGATCCAAATATGGAGATATTAG 802
QY 402 erLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArg--ValValLysLysG 421
Db 803 CACTCATGGAGTGGATPACATTCATATGATATATATGGCGCGGGTAAACCAAGAGAG 862
QY 421 luLeuLysValGlyProTyrAlaAsn--ThrThrArgTyrSer 434
Db 863 AATTAATTTGGGGGCTTTTGGCAACAACATCAACCAATACCG 904

RESULT 10
CF723353
LOCUS
DEFINITION
IMAGE:30546171 5', mRNA sequence.
EST.
CF723353.1 GI:37597521
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 694)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
```

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Dr. M. Bento Soares, University of Iowa
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers

source

1..694
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30546171"
/tissue_type="whole brain"
/dev_stage="1.5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GVO"
note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
is CGAATCGAAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 3,72e-129 Length: 694
Score: 1142.00 Matches: 225
Percent Similarity: 98.27% Conservative: 2
Best Local Similarity: 97.40% Mismatches: 3
Query Match: 40.14% Indels: 2
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x CF723353 (1-694)

QY 78 AsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHis 97
Db 4 AACCTCTTGCTTGGACTTTGACCGTGTACACGACAGAGAGATCTACGATGACCAT 63
QY 98 ArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerPro 117
Db 64 CGCAAGTTTCACTTCGGATCTTATATGACAGGAGGAGGCCACCTCTGCTACCT 123
QY 118 SerSerArgLeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyI 137
Db 124 AGCAGAGAGCTGATGTTGTTAATGTGACTACTCTCCCTGGGAGGTCACTTCTGGAAT 183
QY 137 eGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerAr 157
Db 184 CCAAGGGGCGCATCATCTCTGAGAGATGGAATATGATGAGGAGGCGGCATCATCCCG 243
QY 157 gIlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLe 177
Db 244 GATCTTTGAGAGCGGAAAAATGTGAGGTACACGTACTTAGAGAGTCCATGCTGCTTCA 303
QY 177 uLeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerVa 197
Db 304 TCTCCACGCCAGGCGCATCTTCGAGTTTGACAGAGATGACCGCTCTCTCTGT 363

is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores: 1.84e-125 Length: 758
 Pred. No.: 1112.50 Matches: 194
 Score: 93.20% Conservative: 39
 Percent Similarity: 77.60% Mismatches: 16
 Best Local Similarity: 39.10% Indels: 1
 Query Match: 14 Gaps: 1
 DB:

US-10-029-020-14_COPY_1760_2300 (1-541) x CB723840 (1-758)

QY 291 GlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePhe 310
 DB 9 GGTTCCTCTGTACCATCAGGTACCAAGGTGGGCCCCCTGTGGACAAGCAGATTAC 68
 QY 311 ArgPheThrGluGluGlyMetValAsnAlaArgPheAspTyrAsnTyr---AspAsnSer 329
 DB 69 AGTTTCTCTGAGAGAGGATGATCAACGCCAGGTTTGATTATACCTATCAGCAATAGC 128
 QY 330 PheArgValThrSerMetGlnAlaValIleAsnGluThrProLeuIleAspLeuTyr 349
 DB 129 TTCCGATTCGCCAGCATCAACCCCTCATTTAGCAGACTCCCTCTCTGTGACCTCTAC 188
 QY 350 ArgTyrAspValSerGlyThrGluGlnPheGlyLysPheGlyValIleTyrTyr 369
 DB 189 CGCTATGATGAGATTTCCGGCAAGGTGGAAACATTCGGCAAGTTGGGGTCTACTAC 248
 QY 370 AspIleAsnGlnIleThrAlaValMetThrHisThrLysHisPheAspAlaTyr 389
 DB 249 GACATCAACAGATCATCACCTGCGTCAATGACCTTAGCAAGCACTTTGACACCCAT 308
 QY 390 GlyArgMetLysGluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrVal 409
 DB 309 GGGCGCATCAAGGAAGTGAATATGATGTTCCGGTCCCTCATGATCTGATGATGACTGTG 368
 QY 410 GlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsn 429
 DB 369 CAATATGACAGTATGGTAGGTCATCAAGAGGGAACCTAGAGCCCTATGCCAAC 428
 QY 430 ThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsn 449
 DB 429 ACCACAAGTACACTATGATGATGATGACCGGAGCGCCAGCTCCAGAGGTGGCGGTCAAT 488
 QY 450 AspLysProLeuTyrArgTyrSerTyrAspLeuAsnGlnValAsnLeuHisLeuLeuSerPro 469
 DB 489 GACCGGCTACCTGGCGCTATGCTATGACCTCAATGGGAACCTGACCTTCTTAACCCA 548
 QY 470 GlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArgLeu 489
 DB 549 GGAACAGTGTCTGCTCATGCCCTTACGCTATGACCTCGGTGACCGGAATAACAGGCTA 608
 QY 490 GlyAspValGlnTyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIle 509
 DB 609 GGGGACGTGCAGTACAAAATCGATGACGATGGCTATTTGTGCCAGAGAGGGGTGACATC 668
 QY 510 PheGluTyrAsnSerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSer 529
 DB 669 TTTGAATACAACTCCAAAGGGCTTCTGACAGAGGATCAACAAGGCCGCGGTGGAGC 728
 QY 530 ValArgTyrArgTyrAspGlyLeuGlyArg 539
 DB 729 GTGAGTACCTGATGACGAGGTGGCGGC 758

RESULT 12
 BE741784
 LOCUS
 DEFINITION

757 bp mRNA linear EST 15-SEP-2000
 BE741784
 601595546f1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949571 5',

QY 197 lThrMetProAsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrAR 217
 DB 364 GACCATGCCCAACGTAGCCGCGACAGCTGGAGACCATCCGCTCAGTGGGCTACTACAG 423
 QY 217 gAsnIleTyrGlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspG1 237
 DB 424 GAACATCTACAGCCCGCGAGGCAACGCTCAGTCAATCAGGACTTCACCTGAGGATGG 483
 QY 237 yHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyly 257
 DB 484 ACACCTGTATCATACTCTACTTGGCACCGCGCGGTGATTACAAAGTATGGCAA 543
 QY 257 sIeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspG1 277
 DB 544 GTTGTCAAAGCTGGCGAGACTCTGTATGACACCACTAAGTAAAGCTTCACCTACGACA 603
 QY 277 uThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleAr 297
 DB 604 GACCGCAGGAGTGTGAAGACTGTCACTACAGATGA-GGGTTCACTGCACTATCCG 662
 QY 297 gTyrArgGlnIleGlyProLeuIleAspArg 307
 DB 663 CTACCGTCAGATTGGGCCCTGATTGATCGG 693

RESULT 11
 CB723840
 LOCUS
 DEFINITION

UI-M-FYO-cez-f-24-0-UI.r1 NIH BMAP_FYO Mus musculus cDNA clone
 IMAGE:6844153 5', mRNA sequence.
 CB723840
 CB723840.1 GI:29780982
 EST.
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousef1.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

Location/Qualifiers
 1..758
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6844153"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FYO"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaudo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail

FEATURES

source

http://www.genoscope.cns.fr/tetraodon.

FEATURES

Location/Qualifiers
1..868
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="063M14"
/clone_lib="G"
/note="Genoscope sequence ID : COBG063BG07SP1-end : PUC-ori"

ORIGIN

Alignment Scores:
Pred. No.: 1,06e-123 Length: 868
Score: 1099.00 Matches: 201
Percent Similarity: 90.37% Conservative: 43
Best Local Similarity: 74.44% Mismatches: 26
Query Match: 38.63% Indels: 1
DB: 29 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x CNS03W3H (1-868)

```
QY 172 LysSerMetValLeuLeuLeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsn 191
Db 58 CAGTCAATGGTCTCTCTGCTGCACAGTCAACGGCAGTACATATTTGATTGACTCTCAG 117
QY 192 AspArgLeuSerSerValThrMetProAnValAlaArgGlnThrLeuGluThrIleArg 211
Db 118 GACCGGCTGGCTGGCTGCAATCGGAGTGGCCGTTACACCATGACGACGATCCCGC 177
QY 212 SerValGlyTyrTyrArgAsnIleTyrGlnProGluGlyAsnAlaSerValIleGln 231
Db 178 TCGGTGGGCGYCMYMGAAACCTTACCCCTCCAGAAAGCAACGCTCGGTGGGGTA 237
QY 232 AspPheThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgArgVal 251
Db 238 GATTACAGCGAGGATGGCCCTCTGCTGAGAGTGGCTCACTTGGGCACGGGTGCCCGGTC 297
QY 252 IleTyrLysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysVal 271
Db 298 CTCTACAGTACCGCGCGAGAACAGCTGTGGAGATCTGTACGACGACGACGAGGTC 357
QY 272 SerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGly 291
Db 358 AGTTTACCTACACAGCGCGGAGCTGTGATGAAGCCGTTAACTGACGACGATGAGGC 417
QY 292 PheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArg 311
Db 418 TTCATCTGCTCCATTCGTTATCGCAAGTGGGCCCACTGATCGACGACGAGATTTCCGC 477
QY 312 PheThrGluGluGlyMetValAsnAlaArgPheAspTyrAsnTyrAsnSerPheArg 331
Db 478 TTCAGGAAAGAGCGCATGGTCAATGGCGCTTTGATTACCTACGACGACGACGCTGCGT 537
QY 332 ValThrSerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyr 351
Db 538 GTACACGAGCTGCAAGCGCTCATCAATGAAGAACACCACTGCCTATCGATCTCTATCAT 597
QY 352 AspAspValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrThrAspIle 371
Db 598 GATGACATCTCGGAAAGGTGGAAACGTTTGGAAAGTTTGGAGTCACTATTATGATATC 657
QY 372 AsnGlnIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArg 391
Db 658 AATCAATCATATCCACGGCGATGATGACCTACACCAACACITTTGATGACACGAGGACG 717
QY 392 MetLysGluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrValGlnTyr 411
Db 718 ATCAAGAGAGATCCAGTATGATGATATTCGTTTCGTTTCGTTTCGTTTCGTTTCGTT 777
QY 412 AspAsnMetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnThrThr 431
Db 778 GACGATGTGGGCGAGTGATCAAGAGAGAGATCAAAATC-GGCCCTCTTTTGGAAACACCA 836
```

QY 432 ArgTyrSerTyrGluTyrAspAlaAspGly 441
Db 837 AAGTACGCTWACGAGTACGACGTGACGCGC 866

RESULT 14

CF534506

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CF534506 856 bp mRNA linear EST 12-SEP-2003
UI-M-G10-chg-a-23-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
IMAGE:30536758 5', mRNA sequence.

CF534506 GI:34586474

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 856)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

http://genome.uiowa.edu/distribution/mousefl.html

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers

1..856

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CS7BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:30536758"

/tissue_type="whole brain"

/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"

/lab_host="PH108 (T1 phage resistant)"

/clone_lib="NIH_BMAP_G10"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonato, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is AGCGACAG. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 1,82e-123 Length: 856

Score: 1097.00 Matches: 205

Percent Similarity: 98.57% Conservative: 2

Best Local Similarity: 97.62% Mismatches: 3

Query Match: 38.56% Indels: 0

DB: 14 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x CF534506 (1-856)

QY 332 ValThrSerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyr 351

Db 3 GTGACTAGCATGCGGCTGTGATCAATGAGACCCACTGCGCCATTGACCTCTACCGTAT 62

QY 352 AspAspValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIle 371
Db 63 GATGATGTCAGGAGACAGACAGCTTTGGAGAGTTGGTGTCTACTACTACGACATC 122
QY 372 AsnGlnIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArg 391
Db 123 AACCAAGATCATTTACCACAGCCGTCATGACCCACACCAACCACTTTGATGCTTATGGCAGG 182
QY 392 MetLysGluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrValGlnTyr 411
Db 183 ATGAAGGAAGTACAGTATGAGATTTTCGGTCACTCATGTACTGATGATGCTTCTAGTAT 242
QY 412 AspAsnMetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnThrThr 431
Db 243 GATAACATGGGACGGTAGTGAAGAGGAGCTGAAGGTGGGACCTATGCCAACACTACC 302
QY 432 ArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLys 451
Db 303 CGTACTCTATGATGATGATGCTGATGCCAGCTGCAGACAGTCTCCATCAATGACACAG 362
QY 452 ProLeuTyrArgTyrSerTyrAspIleAsnGlyAsnLeuHisLeuLeuSerProGlyAsn 471
Db 363 CCACCTCTGGCGCTACAGCTATGACCTCAATGGGAACCTACACTTGCTGAGCCCTGGGAC 422
QY 472 SerAlaArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArgLeuGlyAsp 491
Db 423 AGCGACCGCTCACACCACTACGATGATGACCTCCGTCAGCCGATCACCAGGTGGTGTAT 482
QY 492 ValGlnTyrLysMetAspGlyAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGlu 511
Db 483 GTACAGTACAAGATGATGATGATGCTTCTGAGGACGGGTGGGATGCTTCGAG 542
QY 512 TyrAsnSerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTyrSerValArg 531
Db 543 TACAATTCAGCGGGGCTCATCANAGCCATACACCGGCTAGTGGGTGGATGTCAGG 602
QY 532 TyrArgTyrAspGlyLeuGlyArgVal 541
Db 603 TACCGATACGATGGCTGGGACGGAGTA 632
RESULT 15
CNS02PHB 1031 bp DNA linear GSS 01-SEP-2000
LOCUS Tetraodon nigroviridis genome survey, sequence T7 end of clone
DEFINITION 155F06 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL207992
VERSION AL207992.1 GI:7866811
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1
Roest Crolius.H., Jaillon.O., Dasilva.C., Bouneau.L., Fisher.C.,
Bernot.A., Fizames.C., Wincker.P., Brottier.P., Quetier.F.,
Saurin.W. and Weissenbach.J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
20296633
PUBMED 10835645
REFERENCE 2
AUTHORS Roest Crolius.H., Jaillon.O., Dasilva.C., Ozouf-Costaz.C.,
Fizames.C., Fischer.C., Bouneau.L., Billault.A., Quetier.F.,
Saurin.W., Bernot.A. and Weissenbach.J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837

PUBMED 10899143
REFERENCE 3 (bases 1 to 1031)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
location/Qualifiers
source
1..1031
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="155F06"
/clone_lib="G"
/note="Genoscope sequence ID : COAG155DC03LPI-end : T7"
ORIGIN
Alignment Scores:
Pred. No.: 9 27e-120 Length: 1031
Score: 1068.00 Matches: 202
Percent Similarity: 92.98% Conservative: 10
Best Local Similarity: 88.60% Mismatches: 16
Query Match: 37.54% Indels: 1
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1760_2300 (1-541) x CNS02PHB (1-1031)
QY 314 GluGluGlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThr 333
Db 11 GAAGAGGGATGGTCAC -GCCAGATTAGACTATCTCTATGACACAGTTTCAGAGTTACA 69
QY 334 SerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAsp 353
Db 70 AGCATGTAGCBGTGATCAATGAACCTCCACTGCCATCGACTTGTATCGTACGACGAC 129
QY 354 ValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGln 373
Db 130 GTGCTGGCAAAACAGAGCAGCTTTGGTAAATTTGGAGTCATCTATTATGACATAAACCAA 189
QY 374 IleIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLys 393
Db 190 ATTATTACAACAGCCGTCATGACCCACACACCTTCCATGCTTATGATGCGCGGTGAAG 249
QY 394 GluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsn 413
Db 250 GAGTCCCAATATGAATATTTCCGATCTCTTATGTTCTGGATGATGGTGCAGTATGACAAC 309
QY 414 MetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyr 433
Db 310 ATGGGGCGTGTGTGGCCAGAGCTAAAGTGGACCCGTCATGCCAACACAGACGCTAT 369
QY 434 SerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeu 453
Db 370 ACATATGAATATGATGCCATGACCACTTCAGTAGTCTCCATCAATGACAGCCCCCTG 429
QY 454 TrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAla 473
Db 430 TGGAGGTACAGCTATGACCTAAATGGAACCTCCACCTTCTCAGCCCTGGAATATGTCGA 489
QY 474 ArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGln 493
Db 490 CGCTCACCCCGCTACGGTATGATGAACAGGACCCGATCTACTCGCTTGGCGGATGACAG 549
QY 494 TyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsn 513
Db 550 TACAGGATGACAGGACGGTTCCTCAAGCAGCAGGAGGAATGACTACTCTCGAATACMAC 609
QY 514 SerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTyrSerValArgTyrArg 533

Mon Aug 16 09:01:11 2004

Db 610 TCAGCAGCCTTCTAGTTAGGTGTACAAAGTGACGGTTGAGCATACATATCGG 669

Qy 534 TyrAspGlyLeuGlyArgVal 541

Db 670 TATGATGGCTGGGAAGGAGGGTG 693

Search completed: August 14, 2004, 18:04:37
Job time : 6395.74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 917.682 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: / US-10-029-020-14_COPY_1760_2300

Perfect score: 2845

Sequence: 1 YYIGADGSLRLLLANGMEVA.....YNRAGSWSVRYVDGLGRV 541

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spo1/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N_Geneseq_29Jan04 -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CGN_1.1.1868 @runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO.MWAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|------------|--------------------|
| 1 | 2845 | 100.0 | 8354 | 6 ABS52100 | ABS52100 Human TEN |
| 2 | 2845 | 100.0 | 8645 | 6 ABS78652 | ABS78652 Human CDN |
| 3 | 2821.5 | 99.2 | 8438 | 6 ABN85378 | ABN85378 Human NOV |
| 4 | 2233 | 78.5 | 3614 | 4 AAH14183 | AAH14183 Human CDN |
| 5 | 2233 | 78.5 | 8473 | 6 ABQ82345 | ABQ82345 Human NOV |
| 6 | 2233 | 78.5 | 8487 | 6 ABQ82346 | ABQ82346 Human NOV |
| 7 | 2233 | 78.5 | 8645 | 6 ABQ82344 | ABQ82344 Human NOV |
| 8 | 2233 | 78.5 | 8675 | 6 ABQ82343 | ABQ82343 Human NOV |

| | | | | | |
|----|--------|------|-------|------------|--------------------|
| 9 | 2159 | 75.9 | 3270 | 4 AAH14671 | AAH14671 Human CDN |
| 10 | 2136.5 | 75.1 | 9058 | 7 ACC72051 | ACC72051 BCU0205A |
| 11 | 2136.5 | 75.1 | 9695 | 7 ACC72052 | ACC72052 BCU0205B |
| 12 | 2136.5 | 75.1 | 9729 | 5 AAS14089 | AAS14089 Human FCT |
| 13 | 2136.5 | 75.1 | 9729 | 5 ADB32028 | ADB32028 Human FCT |
| 14 | 2136.5 | 75.1 | 9826 | 5 AAS14085 | AAS14085 Human FCT |
| 15 | 2136.5 | 75.1 | 9826 | 5 ADB32023 | ADB32023 Human FCT |
| 16 | 2022.5 | 72.1 | 13202 | 4 AAK51828 | AAK51828 Human pol |
| 17 | 2048.5 | 72.0 | 12879 | 6 AAK92230 | AAK92230 Prostate |
| 18 | 1744.5 | 61.3 | 3312 | 4 AAK52812 | AAK52812 Human pol |
| 19 | 1254.5 | 44.1 | 1227 | 5 AAS68861 | AAS68861 DNA encod |
| 20 | 999 | 35.1 | 812 | 4 AAH06979 | AAH06979 Human CDN |
| 21 | 982.5 | 34.5 | 806 | 4 AAI23889 | AAI23889 Probe #13 |
| 22 | 982.5 | 34.5 | 806 | 4 ABA69008 | ABA69008 Human foe |
| 23 | 982.5 | 34.5 | 806 | 4 AAI49198 | AAI49198 Probe #17 |
| 24 | 982.5 | 34.5 | 806 | 4 ABA51018 | ABA51018 Human bre |
| 25 | 982.5 | 34.5 | 806 | 4 ABA35953 | ABA35953 Probe #14 |
| 26 | 982.5 | 34.5 | 806 | 4 AAK43120 | AAK43120 Human bon |
| 27 | 982.5 | 34.5 | 806 | 4 AAK17327 | AAK17327 Human bra |
| 28 | 982.5 | 34.5 | 806 | 4 ABA42754 | ABA42754 Human liv |
| 29 | 982.5 | 34.5 | 806 | 5 AAI09491 | AAI09491 Probe #94 |
| 30 | 982.5 | 34.5 | 806 | 6 ABS17195 | ABS17195 Human gen |
| 31 | 982.5 | 34.5 | 1973 | 4 AAI14695 | AAI14695 Probe #46 |
| 32 | 982.5 | 34.5 | 1973 | 4 ABA56427 | ABA56427 Human foe |
| 33 | 982.5 | 34.5 | 1973 | 4 AAI36063 | AAI36063 Probe #47 |
| 34 | 982.5 | 34.5 | 1973 | 4 ABA45900 | ABA45900 Human bre |
| 35 | 982.5 | 34.5 | 1973 | 4 ABA26060 | ABA26060 Probe #45 |
| 36 | 982.5 | 34.5 | 1973 | 4 AAK30099 | AAK30099 Human bon |
| 37 | 982.5 | 34.5 | 1973 | 4 AAK04592 | AAK04592 Human bra |
| 38 | 982.5 | 34.5 | 1973 | 4 ABS29749 | ABS29749 Human liv |
| 39 | 982.5 | 34.5 | 1973 | 5 AAI04501 | AAI04501 Probe #44 |
| 40 | 982.5 | 34.5 | 1973 | 6 ABS04679 | ABS04679 Human gen |
| 41 | 882.5 | 31.0 | 640 | 4 AAL26459 | AAL26459 Human bre |
| 42 | 764 | 26.9 | 843 | 4 AAH05415 | AAH05415 Human CDN |
| 43 | 764 | 26.9 | 2157 | 4 AAH14096 | AAH14096 Human CDN |
| 44 | 607.5 | 21.4 | 487 | 4 AAL16510 | AAL16510 Human bre |
| 45 | 585.5 | 20.6 | 410 | 5 AAS68859 | AAS68859 DNA encod |

ALIGNMENTS

RESULT 1

ABS52100

ID ABS52100 standard; DNA; 8354 BP.

XX ABS52100;

XX ABS52100;

DT 05-NOV-2002 (first entry)

XX Human TEN-M4-like gene.

DE

XX

XX

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

```
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX WO200257453-A2.
XX
XX
XX
XX
XX 25-JUL-2002..
XX
XX 19-DEC-2001; 2001WO-US050331.
XX
XX 19-DEC-2000; 2000US-0256704P.
XX 20-DEC-2000; 2000US-0257314P.
XX 02-MAY-2001; 2001US-0288153P.
XX 29-MAY-2001; 2001US-0294075P.
XX 24-JUL-2001; 2001US-0307506P.
XX 10-AUG-2001; 2001US-0311590P.
XX 10-AUG-2001; 2001US-0311613P.
XX 29-AUG-2001; 2001US-0315617P.
XX 14-SEP-2001; 2001US-0322358P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
XX Stone DJ, Anderson D, Shinkets RA, Burgess CE, Zerhusen BD, Liu X;
XX Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX WPI; 2002-590744/63.
XX P-PSDB; ABG70388.
XX
XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
XX atherosclerosis, metabolic disorders, diabetes, obesity, infectious
XX disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
XX cancer.
XX
XX Claim 8; Page 50-52; 318pp; English.
XX
XX The present invention relates to new NOVX polypeptides. The invention is
XX useful for treating or preventing a NOVX-associated disorder such as
XX cardiomyopathy or atherosclerosis, where the disorder is related to cell
XX signal processing and metabolic pathway modulation in a subject,
XX preferably human. The invention is also useful for treating metabolic
XX disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
XX disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
XX disorders, haematopoietic disorders and various cancers. The molecules of
XX the invention are also useful for treating or preventing cirrhosis,
XX pancreatitis, learning and memory defects, infertility, congenital heart
XX defects, acne, hair growth, pigmentation disorders, reproductive health,
XX respiratory disease, gastro-intestinal diseases, endocrine disorders,
XX neurological diseases, bone marrow transplantation, urinary system
XX allergy and inflammation, nephrological disorders, age-related
XX disorders, neuropsychiatric disorders and age-related disorders. The
XX present nucleic acid sequence represents a NOVX gene. This sequence
XX encodes a NOVX protein of the invention
XX
XX Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 2,79e-278 Length: 8354
XX Score: 2845.00 Matches: 541
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-10-029-020-14_COPY_1760_2300 (1-541) x ABS52100 (1-8354)
XX
XX 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
XX
XX 5312 TACTACATCGGGCGCGATGGCTCCCTTGGCGGTGCTGCTGGCCACGCGATGAGGTGGCG 5371
XX
```


| | | | | | | | |
|----|------|--|------|----------------------------------|---|--|------|
| QY | 81 | SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe | 100 | Db | 6681 | GGCCAGCTGCGACAGTCTCCATCAATGACAGCCACTCTGGCGCTACAGCTACGACCTC | 6740 |
| Db | 5601 | TCTCTGCACTTGTATCGCGTAAACAGCAGAGAGATCTATGATGACCCAGCGAGTTC | 5660 | QY | 461 | AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr | 480 |
| QY | 101 | ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerArg | 120 | Db | 6741 | AATGGAACTCTGACCTTACTGAGCCTGGAAACAGTGCACGGCTCACACCACTACGGTAT | 6800 |
| Db | 5661 | ACCCCTCGGATTTCTGTACACAGCGCGGGCGCCAGCCTCTGGTCCACAGCAGCAGG | 5720 | QY | 481 | AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrIleMetAspGluAspGly | 500 |
| QY | 121 | LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly | 140 | Db | 6801 | GACATCCGCGACCGCATCACTCGGTGGGTGACGTGCAATACAAGATGGATGAGGATGC | 6860 |
| Db | 5721 | CTGAATGGTGTCAACGTGACATCTCCCTCGGGGTTACATGCTGGCATCCAGAGGGC | 5780 | QY | 501 | PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys | 520 |
| QY | 141 | IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla | 160 | Db | 6861 | TTCTCTGAGCAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGCTGCTCATCAAG | 6920 |
| Db | 5781 | ATCATGCTCTGAAGAATGAATACGACAGCGGGCGCCATCATCATCCAGATCTTCGCT | 5840 | QY | 521 | AlaTyrAsnArgAlaGlySerTyrPheValArgTyrArgTyrAspGlyLeuGlyArg | 540 |
| QY | 161 | AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer | 180 | Db | 6921 | GCTTACACCGGCTGGAGCTGGAGTGTGAGTACCGCTACGATGGCTGGCGCGCGC | 6980 |
| Db | 5841 | GATGGGAACATGGAGCTACATCTTACATGAGAGTCCATGCTGCTTACTACACAGC | 5900 | QY | 541 | Val 541 | |
| QY | 181 | GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro | 200 | Db | 6981 | GTG 6983 | |
| Db | 5901 | CAGAGCAGTATATCTTTGAGTTCGACAGATGACCGCTCTCTTCTGTGACATGCC | 5960 | RESULT 3 | | | |
| QY | 201 | AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr | 220 | ABN85378 | | | |
| Db | 5961 | AACGTGGCGGCGAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAATCTAT | 6020 | ABN85378 standard; DNA; 8438 BP. | | | |
| QY | 221 | GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu | 240 | XX | 21-OCT-2002 | (first entry) | |
| Db | 6021 | CAGCCCTCGAGGCAATGCTCAGTATACAGGACTTCACTGAGGATGGCACCTCCT | 6080 | AC | Human NOV1, TEN-M4 like protein, coding sequence. | | |
| QY | 241 | HisThrPheTyrLeuGlyThrArgValIleTyrLysTyrGlyLysLeuSerIys | 260 | XX | Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive; | | |
| Db | 6081 | CACACCTTCTACCTGGCGCACTGGCGCGAGGTGATATACAGTATGGCAACTGTCAAG | 6140 | KW | Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic; | | |
| QY | 261 | LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly | 280 | KW | Antibacterial; Nephrotoxic; Hepatotropic; Neuroprotective; Nootropic; | | |
| Db | 6141 | CTGCGACAGACGCTCTATGACACCAACAGGTGAGTTCCTATGACGAGCGGCGAGC | 6200 | KW | Antitubercular; Virucide; Antiparasitic; Relaxant; Anticonvulsant; | | |
| QY | 281 | MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln | 300 | KW | Gene Therapy; NOV; cancer; heart disease; inflammation; | | |
| Db | 6201 | ATGCTGAAGACCATCAACCTACAGATGAGGCTTCACTGCACCATCCGCTACCGTCAG | 6260 | KW | autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity; | | |
| QY | 301 | IleGlyProLeuLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla | 320 | KW | asthma; IgA nephropathy; cirrhosis; arthritis; Alzheimer's disease; | | |
| Db | 6261 | ATTGGGCCCCGCTGATTGACCGACAGATCTTCGCTTCACTGAGGAGGATGGTCAACGCC | 6320 | KW | infection; stroke; muscular dystrophy; epilepsy; wasting disorder; | | |
| QY | 321 | ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn | 340 | XX | TEN-M4 like protein; chromosome 11; gene; ds. | | |
| Db | 6321 | CGTTTGTACTCAACTATGACAAACAGCTTCGGGTGACGATGCGGCTGTGATCAAC | 6380 | OS | Homo sapiens. | | |
| QY | 341 | GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln | 360 | Key | Location/Qualifiers | | |
| Db | 6381 | GAGACCCCACTGCCATTTGATCTCTATCGCTATGATGATGTGTGAGCAAGACAGCAG | 6440 | Key | 4..8395 | | |
| QY | 361 | PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet | 380 | CDS | /*tag= a | | |
| Db | 6441 | TTTGGGAAGTTTGGTGTCTTACTATGACATTAACACCATGATCATCACACAGCTGTATG | 6500 | FT | /trans_except= (pos: 1138..1147,aa:Met) | | |
| QY | 381 | ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe | 400 | FT | /product= "NOV1 protein" | | |
| Db | 6501 | ACCCACACAGCATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 6560 | PN | WO200255704-A2. | | |
| QY | 401 | ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValLysLys | 420 | PD | 18-JUL-2002. | | |
| Db | 6561 | CGCTCGCTCATGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 6620 | PF | 09-JAN-2002; 2002WO-US000554. | | |
| QY | 421 | GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp | 440 | PR | 09-JAN-2001; 2001US-0260417P. | | |
| Db | 6621 | GAGCTGAGGTAGGACCCCTACGCCAATACCTCGCTCTCTCTCTCTCTCTCTCTCTCT | 6680 | PR | 10-JAN-2001; 2001US-0260831P. | | |
| QY | 441 | GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu | 460 | PR | 28-FEB-2001; 2001US-0272338P. | | |
| | | | | PR | 09-MAR-2001; 2001US-0274876P. | | |
| | | | | PR | 18-APR-2001; 2001US-0284704P. | | |
| | | | | XX | (CURA-) CURAGEN CORP. | | |
| | | | | XX | Padigar M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA; | | |
| | | | | PI | Zhong M, Gangoli EA, Burgess CE, Patturajan M, Vernet CAM; | | |
| | | | | PI | Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM; | | |
| | | | | PI | Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K; | | |
| | | | | PI | MacDougall J, Malyankar U, Millet I, Peyman J, Smithson G; | | |
| | | | | PI | Gunther E, Stone DJ; | | |
| | | | | XX | WPI; 2002-590674/63. | | |

XX AAH14183;
AC
XX
XX
DT
XX
XX
DE Human cDNA sequence SEQ ID NO:11430.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX
XX EP1074617-A2.
PN
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
PR
XX 27-AUG-1999; 99JP-00300253.
PR
XX 11-JAN-2000; 2000JP-00118776.
PR
XX 02-MAY-2000; 2000JP-00183767.
PR
XX 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
DR
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
PT
XX
XX Claim 8; SEQ ID NO 11430; 2537pp + Sequence Listing; English.
PS
XX
XX The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the polynucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the polynucleotide comprises at least 15 nucleotides and the combination of oligonucleotide comprises at least 15 nucleotides and those defined in the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX
XX Sequence 3614 BP; 967 A; 898 C; 928 G; 821 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 1.56e-216 Length: 3614
Score: 2233.00 Matches: 404
Percent Similarity: 88.91% Conservative: 77
Best Local Similarity: 74.68% Mismatches: 60
Query Match: 78.49% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AAH14183 (1-3614)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20

Db
272 TACCAGATTGGTTATGACGGCTCCCTCAGAAATATCTACGCCAGTGGCTGGACTCACAC 331
QY
21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db
332 TACCAAAACAGAGCCGACGCTTCTGGCTGGCACCCTATCCGACGGTTCGCCAAAGAAC 391
QY
41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrPargGlnArgLysGluGln 60
Db
392 ATGACTTTGGCTGGCGAGAACGGTCAAAACTTTGGTGAATGGAGATTCCGAAAAGAGCAA 451
QY
61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db
452 GCCCAAGGGGAAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCGAAACCTCCCT 511
QY
81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db
512 TCAGTTGACTTTTATCGAACAACAAGACAGAAAGATCTATGACGACCACCGTAAATTT 571
QY
101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrPargProSerSerArg 120
Db
572 CTACTGAGGATCGCTTACGACACGCTCGGCGACCCGACTCTCTGGCTGCCAAGCAGCAAG 631
QY
121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db
632 CTGATGGCGCTCAATGTCTACCTATTCATCCACAGGTCAAATGCCAGATCCACGCGAGGC 691
QY
141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db
692 ACCACTAGCGAGAAAGTAGATTATGACGACGGGAGGATCGTGCTCGGCTTTGCT 751
QY
161 AspGlyLysThrTyrPargTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db
752 GATGGTAAACATGAGTTACACATATTTAGAAAAGTCCATGGTTCTCTGCTTCATAGC 811
QY
181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db
812 CAGCGGAGTACATCTCGAATACGATATGTGGACCGCTCTGCTGCATCACCATGCC 871
QY
201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db
872 AGTGTGGCTGGCCACACCATGACGACCATCCGATCCATGGTACTACTACCGCAACATATAC 931
QY
221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db
932 AACCCCGGAAAGCAACGGCTCCATCATCCGAGTACACGAGGAGGCTGCTCTCTA 991
QY
241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db
992 CAACACAGCTTTCTTGGGTACAAAGTCGGAGGCTCTTATCAATACAGAGGCGACTAGG 1051
QY
261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db
1052 CTCTCAGAAATTTATATGATAGCACAAAGTCAAGTTCAGTTTACCTATGATGAACAGCAGGA 1111
QY
281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db
1112 GTCCTAAAGACAGTAAACCTCCAGAGTGATGGTTTATTTTCACCACTTAGATACAGGCA 1171
QY
301 IleGlyProIleLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db
1172 ATGGTCCCTGATTGACGACGAGATTTTCGCTTTAGTGAAGATGGGTGTAATGCA 1231
QY
321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db
1232 AGATTGACTATGATGACAAACAGCTTTCCAGTGACAGCAGTACAGGGTGTGATCAAT 1291
QY
341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db
1292 GAAACGCCACTGCCTTATGATCTGTATGATGATGATTTCTGGCAAAAGTTCAGCAG 1351
QY
361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleThrThrAlaValMet 380

| | | | | | | | |
|----|------|--|------|----------|--|---|------|
| QY | 21 | LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn | 40 | QY | 381 | ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluLeuPhe | 400 |
| DB | 5190 | TACCAAAACAGAGCCGACGCTCTGGCTGGCACCGCTTAATCCGACGGTTGCCAAAAGAAAC | 5249 | DB | 6270 | ACCTATACGAAGCACTTTGATGCTCATCGCCGCTATCAAGAGATTCAATATGAGATATTC | 6329 |
| QY | 41 | ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln | 60 | QY | 401 | ArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValLysLys | 420 |
| DB | 5250 | ATGACTTTGCCCTGGCGAGAACCGGTCAAAACTTGGTGGAAATGAGATTCCGAAAAGAGCAA | 5309 | DB | 6330 | AGGTGGCTCATGACTTGGATTACAATTCAGTATGATTAACATGGTCCGGTACCAAGAGA | 6389 |
| QY | 61 | AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu | 80 | QY | 421 | GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp | 440 |
| DB | 5310 | GCCCAAGGGAAGTCAATGCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCCTT | 5369 | DB | 6390 | GAGATTAAATAGGCCCCCTTCCCAACACCCAAATATGCTTATGATATGATGTTGAT | 6449 |
| QY | 81 | SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe | 100 | QY | 441 | GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu | 460 |
| DB | 5370 | TCAGTTGACTTTGATCGAACAACAAAGACAGAAAAGATCTATGACACACCCGTAATTT | 5429 | DB | 6450 | GCACAGCTCCAAACAGTTTACCTCAATGATAAGATAATCTGGCGGTACCACTACGATCTG | 6509 |
| QY | 101 | ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrPheSerProSerArg | 120 | QY | 461 | AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr | 480 |
| DB | 5430 | CTACTGAGGATCCCTTACGACAGCTCTGGCCACCCGACTCTCTGGCTGCCAAGCAGCAAG | 5489 | DB | 6510 | AATGGAAACCTCCATTTACTGAACCCAGTAACAGTCCGGCTGACACCCCTTCCTAT | 6569 |
| QY | 121 | LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly | 140 | QY | 481 | AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly | 500 |
| DB | 5490 | CTGATGCCCTCAATGTCATCTATCCACAGGTCAAATTCGACGATCCAGCGAGGC | 5549 | DB | 6570 | GACCTCGAGACAGATCACTCGACTGGGTGATGTTCAATATCGTTGGATGAAGATGGT | 6629 |
| QY | 141 | IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla | 160 | QY | 501 | PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys | 520 |
| DB | 5550 | ACCACCTAGCCAGAAAGTAGATTATGACGACAGGGAGGATCGTCTCGGGCTTTGCT | 5609 | DB | 6630 | TTCTCTACGTCAAAGGGGACGGAATCTTTGAATATAGCTCCAAAGGGCTTCACTCGA | 6689 |
| QY | 161 | AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer | 180 | QY | 521 | AlaTyrAsnArgAlaGlySerTyrValArgTyrArgTyrAspGlyLeuGlyArgArg | 540 |
| DB | 5610 | GATGGTAAACATGGAGTTACATATTTAGAAAAGTCCATGTTCTTCTGCTTCATAGC | 5669 | DB | 6690 | GTTTACAGTAAGGCGAGTGGCTGGACAGTCTACCGTTATGACGGCTGGGAAGCGGT | 6749 |
| QY | 181 | GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro | 200 | QY | 541 | Val 541 | |
| DB | 5670 | CAGCGGAGTACATCTTCGAATACGATATGTTGGGACCGCTCTCTGCCATCACCATCCC | 5729 | DB | 6750 | GTT 6752 | |
| QY | 201 | AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr | 220 | RESULT 6 | | | |
| DB | 5730 | AGTGTGGCTGGCCACACCATGACACCATCCGATCCATTTGGCTACTACCGCAACATATAC | 5789 | ABQ82346 | | | |
| QY | 221 | GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu | 240 | ID | ABQ82346 | standard; cDNA; 8487 BP. | |
| DB | 5790 | AACCCCCGGAAACGACCGCTCCATCATCGAGACTACACGAGGAGGCGCTCTCTTA | 5849 | XX | ABQ82346; | | |
| QY | 241 | HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys | 260 | AC | ABQ82346; | | |
| DB | 5850 | CAACAGCTTCTTTGGGTACAAAGTCGAGGGCTTTATTCAAATACAGAAGCGACACTAGG | 5909 | DT | 17-DEC-2002 | (first entry) | |
| QY | 261 | LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly | 280 | DE | Human | NOV15d encoding cDNA SEQ ID NO:41. | |
| DB | 5910 | CTCTCAGAAATTTATATGATAGACACAGAGTCACTTATGATGATGAACACAGCAGA | 5969 | XX | Human; | NOVX; cytototoxic; neuroprotective; anticonvulsant; cardiovascular; | |
| QY | 281 | MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln | 300 | KW | Human; | NOVX; cytototoxic; neuroprotective; anticonvulsant; cardiovascular; | |
| DB | 5970 | GTCCTTAAGACAGTAACCTCCAGAGTCAATGTTTATTTGCACCATTAGATACAGGCAA | 6029 | KW | cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; | | |
| QY | 301 | IleGlyProLeuLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla | 320 | KW | antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; | | |
| DB | 6030 | ATTGCTCCCTCAATTGACAGGACAGATTTTCCGCTTTAGTGAAGATGGGTGAATGCA | 6089 | KW | antianaemic; antibacterial; protozoicide; antihelminthic; gene therapy; | | |
| QY | 321 | ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn | 340 | KW | cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; | | |
| DB | 6090 | AGATTGACTATGAGCTATGACAAACAGCTTTCAGTGACAGATCCAGGGTGTGATCAAT | 6149 | KW | stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; | | |
| QY | 341 | GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln | 360 | KW | Pick's disease; vesicular transport disease; cystic fibrosis; goitre; | | |
| DB | 6150 | GAAACGCCACTCCCTATTGATCTGATCATCAGTTTGTATGATGATCCAGGAGTTCAGGAG | 6209 | KW | diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; | | |
| QY | 361 | PheGlyLysPheGlyValIleTyrAspIleAsnGlnIleThrThrAlaValMet | 380 | KW | ulcerative colitis; gastric disorder; duodenal disorder; infection; | | |
| DB | 6210 | TTTGGAAAGTTTGGAGTTATATATATGATATTAACCCAGATCATTTCTACAGCTGTATG | 6269 | KW | autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; | | |
| | | | | XX | rheumatoid arthritis; gene; chromosome 4; ss. | | |
| | | | | OS | Homo sapiens. | | |
| | | | | XX | | | |
| | | | | XX | Key | Location/Qualifiers | |
| | | | | FT | CDS | 299..8140 | |
| | | | | FT | | /*tag= a | |
| | | | | FT | | /product= "NOV15d" | |
| | | | | XX | | | |
| | | | | XX | WO200262999-A2. | | |
| | | | | XX | | | |
| | | | | PD | 15-AUG-2002. | | |
| | | | | XX | | | |
| | | | | XX | 31-DEC-2001; 2001WO-US049976. | | |
| | | | | XX | | | |
| | | | | XX | 29-DEC-2000; 2000US-0259928P. | | |
| | | | | PR | 02-JAN-2001; 2001US-0259415P. | | |
| | | | | PR | | | |

6266 ACCTATACGACCACTTTGATGCTCATGCCGTATCAAGGAGTTCATATGAGATATTC 6325
401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVallylsys 420
6326 AGGTGCTCATGTACTGGATTACAATTAGTATGATTAACATGGTTCGGGTAAACAGAGA 6385
421 GluLeuValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaasp 440
6386 GAGATTAAATAGGCGCCCTTTCCCAACACCCACCAATATGCTTATGAATATGATGTGAT 6445
441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
6446 GGACAGCTCCAAACAGTTTACTCAATGAAGATAATGTGGCGGTACAACTACGATCTG 6505
461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
6506 AATGAAACCTCATTTACTGAACCCACAGTACGTCGCGTCTGACACCCCTTCGCTAT 6565
481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
6566 GACCTGCGAGACAGATCACTCGACTGGGTGATTTCAATATCGGTGGATGAGATGCT 6625
501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys 520
6626 TTCTTACGTCNAAGSGGACGGAATCTTTGATATAGCTCCAAAGGGCTTCTAACTCGA 6685
521 AlaTyrAsnAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
6686 GTTTACAGTAAGGCAGTGGCTGCAGCAGTATACCTGTTATGACGGCCTGGGAGGCGT 6745
541 Val 541
6746 GTT 6748
RESULT 7
ABQ82344
ID ABQ82344 standard; cDNA; 8645 BP.
XX AC ABQ82344;
XX AC ABQ82344;
DT 17-DEC-2002 (first entry)
XX Human NOV15b encoding cDNA SEQ ID NO:37.
XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
ulcerative colitis; gastric disorder; duodenal disorder; infection;
autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
rheumatoid arthritis; gene; chromosome 4; ss.

09-MAR-2001; 2001US-0279863P.
29-MAR-2001; 2001US-0279832P.
29-MAR-2001; 2001US-0279833P.
13-APR-2001; 2001US-0283889P.
18-APR-2001; 2001US-0284447P.
25-APR-2001; 2001US-0286683P.
29-MAY-2001; 2001US-0294080P.
16-AUG-2001; 2001US-0312915P.
17-AUG-2001; 2001US-031325P.
17-SEP-2001; 2001US-0322699P.
26-NOV-2001; 2001US-0333350P.
(CURA-) CURAGEN CORP.
Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
Rekuda R, Paturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
Gunther E, Smithson G, Millet I, Macdougall JR;
WPI; 2002-732706/79.
P-PSDB; ABP53587.
New NOVX polypeptides and polynucleotides useful for treating NOVX-
associated disorders, such as cancers, neurological disorders, disorders
of vesicular transport, gastrointestinal disorders, and autoimmune
diseases.
Claim 8; Page 114-117; 444pp; English.
The present invention describes novel human proteins designated NOVX,
where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
antidiabetic, antiinflammatory, antirheumatic, antiallergic, virucide,
immunosuppressive, antianaemic, antianaemic, antibacterial, fungicide,
protozoacide and antihelminthic activities, and can be used in gene
therapy. The NOVX proteins, nucleotides or antibodies can be used in the
manufacture of a medicament for treating a syndrome associated with a
human disease selected from NOVX-associated disorder, such as cancers
(e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
disease), disorders of vesicular transport (e.g. cystic fibrosis,
diabetes mellitus, Grave's disease, or goitre), gastrointestinal
disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
and protozoal infections. The NOVX proteins can be used as immunogens to
produce antibodies and as vaccines. The NOVX nucleotide sequences may be
used in chromosome mapping, identifying individuals from minute
biological samples (tissue typing), and in forensic identification of a
biological sample. The present sequence encodes human NOV15b, which is
located on chromosome 4
SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 5,12e-216 Length: 8645
Score: 2233.00 Matches: 404
Percent Similarity: 88.91% Conservative: 77
Best Local Similarity: 74.68% Mismatches: 60
Query Match: 78.49% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1760_2300 (1-541) x ABQ82344 (1-8645)
QY 1 TTTTyrileGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
DB 5302 TACCAGATTGTTATGACGGCTCCCTCAGATTATCTACGCGCTGGCTGACACAC 5361
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
DB 5362 TACCACAAACAGACGCCGACGTTCTGCTGCCACCGCTATATCCGACGGTTCGCAAAAGAAC 5421

| | | | |
|----|------|---|------|
| QY | 41 | ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgGlySerGluCln | 60 |
| Db | 5422 | ATGACTTTTCCCTGGCGAGAACGGTCAAAACTTTGGTGGAAATGGAGATCCGAAAGAGCAA | 5481 |
| QY | 61 | AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu | 80 |
| Db | 5482 | GCCCAAGGAAAGTCAATGCTTTGGCCGCAAGCTCAGGCTTAATGGCAGAAACCTCCT | 5541 |
| QY | 81 | SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgGlyPhe | 100 |
| Db | 5542 | TCAGTTGACTTTTGATCGAACAACAAGACAGAAAGATCTATGACGACCCACCGTAAATTT | 5601 |
| QY | 101 | ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrPsrProSerSerArg | 120 |
| Db | 5602 | CTACTGAGATCGCCTACGACACGGCTCGGCACCGACTCTCTGGCTGCCAAGCAGCAAG | 5661 |
| QY | 121 | LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly | 140 |
| Db | 5662 | CTGATGGCCGCTCAATGTTCACCTATTTCACACAGTCAAAATGGCCAGCATCCAGCAGGC | 5721 |
| QY | 141 | IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla | 160 |
| Db | 5722 | ACCCTACGGAAGTAGATTATGACGACGAGGGAGATCGGTCTCGGGTCTTTGGCT | 5781 |
| QY | 161 | AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer | 180 |
| Db | 5782 | GATGTAACCAATGAGTTACACATATTTAGAAAGTCCATGGTCTCTTCGTTCATAGC | 5841 |
| QY | 181 | GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro | 200 |
| Db | 5842 | CAGCGGCAGTACATCTTCGAATACGATATGTGGACCGCCTGTCTGCCATCCACCATGCC | 5901 |
| QY | 201 | AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr | 220 |
| Db | 5902 | AGTGTGCTGCCACACCATCGACACCATCCGATCCATCGGTACTACTCCGCAACATATAC | 5961 |
| QY | 221 | GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu | 240 |
| Db | 5962 | AAACCCCGGAAAGCAACGCTCCATCATCACGGACTCAACAGGAGAGGGCTGCTTCTA | 6021 |
| QY | 241 | HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys | 260 |
| Db | 6022 | CAACACAGCTTCTTGGGTACAAGTCGAGGGTCTTATTCAAATACAGAGGCGAGACTAGG | 6081 |
| QY | 261 | LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly | 280 |
| Db | 6082 | CTCTCAGAAATTTATATGATGATGACAGAGTCAAGTTCAGTTTACCTATGATGAAACAGCAG | 6141 |
| QY | 281 | MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln | 300 |
| Db | 6142 | GTCTTAAGACAGATAAACCTCCAGAGTGATGTTTATTTCACCCATTAGATACAGGCNA | 6201 |
| QY | 301 | IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla | 320 |
| Db | 6202 | ATTGGTCCCTGATGACAGCGAGATTTTCGGCTTTTAGTCAAGATGGGATGGTAAATGCA | 6261 |
| QY | 321 | ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn | 340 |
| Db | 6262 | AGATTTGACTATAGCTATGACAAACAGCTTTCGAGTGACCAGCATGCGGGTGTGATCAAT | 6321 |
| QY | 341 | GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln | 360 |
| Db | 6322 | GAAACGCCACTGCGCTATTGATCTGTATGATGATGACATTTCTGGCAAAAGTTGAGCAG | 6381 |
| QY | 361 | PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet | 380 |
| Db | 6382 | TTTGAAAGATTTCGAGGTATATATATGATATATAACCATCATTTCTACAGCTGCTAATG | 6441 |
| QY | 381 | ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe | 400 |
| Db | 6442 | ACCTATACAGACATTTGATGCTCATGGCCGTATCAAGGAGATCAATATGAGATATTC | 6501 |

| | | | | |
|----------|---|---|--|------|
| QY | 401 | ArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValValLysLys | 420 | |
| Db | 6502 | AGTGCCTCATGTA | CTGGATTACAATTCAGTATGATAACATGGCTCGGGTAACCAAGAGA | 6561 |
| QY | 421 | GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp | 440 | |
| Db | 6562 | GAGATTAAATAGGCGCCCTTTGGCAACACCAACCAAAATATGCTTATGAATATGATTTGAT | 6621 | |
| QY | 441 | GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu | 460 | |
| Db | 6622 | GGACAGCTCCAAACAGTTTACCTCAATGAAAAGATAATGTGGCGGTACAACCTACGATCTG | 6681 | |
| QY | 461 | AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr | 480 | |
| Db | 6682 | AATGGAAACCTCCATTTACTGAACCCCAAGTAACAGTGGCGTCTGCACCCCTTCGCTAT | 6741 | |
| QY | 481 | AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly | 500 | |
| Db | 6742 | GACCTCGAGACAGAATCATCTCGACTGGGTGATGTTCAATATCGTTGGATGAAGATGGT | 6801 | |
| QY | 501 | PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys | 520 | |
| Db | 6802 | TTCTCAGTCAAGGGGACGGAAATCTTTGAATATAGCTCCAAGGGGCTTTCACTCGA | 6861 | |
| QY | 521 | AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg | 540 | |
| Db | 6862 | GTTTACAGTAAAGGCGAGTGGCTGGACAGTATCTACCGTTATGACGGCTCGGAAGGGCT | 6921 | |
| QY | 541 | Val | 541 | |
| Db | 6922 | GTT | 6924 | |
| RESULT 8 | | | | |
| ABQ82343 | | | | |
| ID | ABQ82343 | standard; cDNA; 8675 BP. | | |
| XX | ABQ82343; | | | |
| XX | 17-DEC-2002 | (first entry) | | |
| XX | Human NOV15a | encoding cDNA SEQ ID NO:35. | | |
| KW | Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; | | | |
| KW | cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; | | | |
| KW | antithematic; antiarthritic; immunosuppressive; antiallergic; virucide; | | | |
| KW | antianemic; antibacterial; protozoacide; antihelminthic; gene therapy; | | | |
| KW | cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; | | | |
| KW | stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; | | | |
| KW | Pick's disease; vesicular transport disease; cystic fibrosis; goitre; | | | |
| KW | diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; | | | |
| KW | ulcerative colitis; gastric disorder; duodenal disorder; infection; | | | |
| KW | autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; | | | |
| KW | rheumatoid arthritis; gene; chromosome 4; ss. | | | |
| XX | Homo sapiens. | | | |
| XX | Key | Location/Qualifiers | | |
| FH | CDS | 151..8328 | | |
| FT | | /*tag= a | | |
| FT | | /product= "NOV15a" | | |
| FT | | /transl_except= (pos:1249..1251,aa:Ser) | | |
| XX | WO200262999-A2. | | | |
| PN | 15-AUG-2002. | | | |
| PD | | | | |
| XX | 31-DEC-2001; 2001WO-US049976. | | | |
| XX | 29-DEC-2000; 2000US-0258928P. | | | |
| PR | 02-JAN-2001; 2001US-0259415P. | | | |
| PR | 04-JAN-2001; 2001US-0259785P. | | | |
| PR | 20-FEB-2001; 2001US-0269814P. | | | |
| PR | 09-MAR-2001; 2001US-0279863P. | | | |

| | | | | | | |
|----|--|------------------|----|------|---|------|
| PR | 29-MAR-2001; | 2001US-0279832P. | QY | 41 | ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluThrArgGlnArgGlyGlnGln | 60 |
| PR | 29-MAR-2001; | 2001US-0279833P. | Db | 5434 | ATGACTTTGCGTGGCAGAACGGTCAAAACCTTGGTGGTAAATGGAGATCCGAAAGAGCAA | 5493 |
| PR | 13-APR-2001; | 2001US-0283889P. | QY | 61 | AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnAspAsnLeuLeu | 80 |
| PR | 18-APR-2001; | 2001US-0284447P. | Db | 5494 | GCCCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAAYCTCCTT | 5553 |
| PR | 25-APR-2001; | 2001US-0286683P. | QY | 81 | SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe | 100 |
| PR | 29-MAY-2001; | 2001US-0294080P. | Db | 5554 | TCAGTTGACTTGTATCGAACACAAAGACAGAAARATCTATGACGACCCAGCTAAATTT | 5613 |
| PR | 16-AUG-2001; | 2001US-0312915P. | QY | 101 | ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerArg | 120 |
| PR | 17-AUG-2001; | 2001US-0313225P. | Db | 5614 | CTACTGAGGATCGCTACGACACGCTCTGGCCACCCGACTCTCTGGTCGCAAGCAAG | 5673 |
| PR | 17-SEP-2001; | 2001US-0322699P. | QY | 121 | LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly | 140 |
| PR | 26-NOV-2001; | 2001US-0333350P. | Db | 5674 | CTGATGGCCGTCATGTCTACCTATTTCATCCACAGGTCAATTTGCCAGCATCCAGCGAGC | 5733 |
| XX | (CURA-) CURAGEN CORP. | | QY | 141 | IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla | 160 |
| XX | Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X; | | Db | 5734 | ACCCTAGCGAGAAAGTAGATTATGACGACGAGCGGAGGAGATCGTGTCTGGGTCTTTGCT | 5793 |
| PI | Malyankar U, Shinkets RA, Tchernev VT, Spaderna SK, Gorman L; | | QY | 161 | AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer | 180 |
| PI | Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S; | | Db | 5794 | GATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGGTCTTCTTCGTTCTAAGC | 5853 |
| PI | Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K; | | QY | 181 | GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro | 200 |
| PI | Gunther E, Smithson G, Millet I, Macdougall JR; | | Db | 5854 | CAGCGCAGTACATCTTCGAATACGATATGTGGACCGCTCTCTGCCATCACCATGCC | 5913 |
| XX | WPI; 2002-732706/79. | | QY | 201 | AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr | 220 |
| XX | P-PSDB; ABP53586. | | Db | 5914 | AGTGTGCTCGCCACACCATGCGACCATCCGATCCATGGCTACTACCGCAACATATAC | 5973 |
| XX | New NOVX polypeptides and polynucleotides useful for treating NOVX- | | QY | 221 | GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu | 240 |
| XX | associated disorders, such as cancers, neurological disorders, disorders | | Db | 5974 | AACCCCGGAAAGCAACGCTCCATCATCAGGACTACCAACGAGGAGGGTGGCTCTTA | 6033 |
| XX | of vesicular transport, gastrointestinal disorders, and autoimmune | | QY | 241 | HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys | 260 |
| XX | diseases. | | Db | 6034 | CAAAAGCTTCTTGGGTACAGTCGAGGGTCTTATTCAAATACAGAGGACAGCTAGG | 6093 |
| XX | Claim 8; Page 110-112; 444pp; English. | | QY | 261 | LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly | 280 |
| XX | The present invention describes novel human proteins designated NOVX, | | Db | 6094 | CTCTCAGAAATTTATATGATAGACACAAAGAGTCAAGTTTACCTATGATGAACAGCAGGA | 6153 |
| XX | where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective, | | QY | 281 | MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln | 300 |
| XX | cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular, | | Db | 6154 | GTCCTAAGACAGATAAACCTCCAGAGTGGTGTATTTATTTGCACCATAGTACAGGCAA | 6213 |
| XX | antidiabetic, antiinflammatory, antineumatic, antiarthritic, virucide, | | QY | 301 | IleGlyProLeuLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla | 320 |
| XX | immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide, | | Db | 6214 | ATTGGTCCCTGATTGACAGCAGATTTTCGCTTTAGTGAAGATGGATGGTAAATGCA | 6273 |
| XX | protozoicide and antihelminthic activities, and can be used in gene | | QY | 321 | ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn | 340 |
| XX | therapy. The NOVX proteins, nucleotides or antibodies can be used in the | | Db | 6274 | AGATTTGACTATAGCTATGACACAGCTTTTCAGTGCACCATGTCAGGGTGTGATCAAT | 6333 |
| XX | manufacture of a medicament for treating a syndrome associated with a | | QY | 341 | GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln | 360 |
| XX | human disease selected from NOVX-associated disorder, such as cancers | | Db | 6334 | GAACCGCCACTCCTATTGATCTGTATCATGTTGATGACATTTCTGGCAAGATTGAGCAG | 6393 |
| XX | (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle, | | QY | 361 | PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet | 380 |
| XX | ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke, | | Db | 6394 | TTTGGAAAGTTTGGAGTTATATATATGATATTAACCATGATCTTTTACAGCTGTAAATG | 6453 |
| XX | ischemic cerebrovascular disease, Alzheimer's disease or Pick's | | QY | 381 | ThrHisThrLysHisPheAspAlaTyrClyArgMetLysGluValGlnTyrGluIlePhe | 400 |
| XX | disease), disorders of vesicular transport (e.g. cystic fibrosis, | | Db | 6454 | ACCTATACGAAGCACTTTGATGCTCATGGCCGATATCAAGGAGATTCAATATGAGATATTC | 6513 |
| XX | diabetes mellitus, Grave's disease, or goitre), gastrointestinal | | QY | 401 | ArgSerLeuMetTyrTyrThrMetThrValGlnTyrAspAsnMetGlyArgValValLysLys | 420 |
| XX | disorders (e.g. ulcerative colitis, or gastric and duodenal disorders), | | | | | |
| XX | autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic | | | | | |
| XX | anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic | | | | | |
| XX | disease), and protozoal infections. The NOVX proteins can be used as immunogens to | | | | | |
| XX | produce antibodies and as vaccines. The NOVX nucleotide sequences may be | | | | | |
| XX | used in chromosome mapping, identifying individuals from minute | | | | | |
| XX | biological samples (tissue typing), and in forensic identification of a | | | | | |
| XX | biological sample. The present sequence encodes human NOV15a, which is | | | | | |
| XX | located on chromosome 4 | | | | | |
| XX | Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other; | | | | | |
| XX | Alignment Scores: | | | | | |
| XX | Pred. No.: 5.14e-216 Length: 8675 | | | | | |
| XX | Score: 2233.00 Matches: 404 | | | | | |
| XX | Percent Similarity: 88.91% Conservative: 77 | | | | | |
| XX | Best Local Similarity: 74.68% Mismatches: 60 | | | | | |
| XX | Query Match: 78.49% Indels: 0 | | | | | |
| XX | DB: 6 Gaps: 0 | | | | | |
| XX | US-10-029-020-14_COPY_1760_2300 (1-541) x ABQ82343 (1-8675) | | | | | |
| QY | 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla | 20 | | | | |
| Db | 5314 TACCAGATTGTTATGACGGTCCCTCAGAAATATCTACGCGAGTGGCGCTGACTCACAC | 5373 | | | | |
| QY | 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn | 40 | | | | |
| Db | 5374 TACCAAAACAGACCGCACGTTCTGCTGGCACCGCTAATCCAGCGTTGCCAAAGAAAC | 5433 | | | | |


```

Db 6514 AGTTCGCTCATGTTACTGGATTACAAATTCAGTATGATACATGGTTCGGGTAACCAAGAGA 6573
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 6574 GAGATTAAATAGGCGCCCTTGGCCAAACACCAACCAATATGCTTATGATATGATGTTGAT 6633
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
Db 6634 GGACAGCTCCAAACAGTTTACCTCAATGAAAGATAATGTGGCGGTACAACTACGATCTG 6693
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 6694 AATGGAAACCTTCATTACTGAACCCCAAGTAACAGTGGCGGTCTGACACCCCTTGGCTAT 6753
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 6754 GACCTCGAGACAGATCACTCGACTGGGTGATGTTCAATATCGGTGGATGAAGATGGT 6813
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys 520
Db 6814 TTCTCTACGTCAAGGGGCGACGGAATCTTTGAATATAGCTCCAAAGGGGCTTCTAACTCGA 6873
QY 521 AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 6874 GTTTACAGTAAGGCGAGTGGCTGGACAGTATCTACCGTTATGACGCGCTGGAGCGGT 6933
QY 541 Val 541
Db 6934 GTT 6936

RESULT 9
AAH14671
ID AAH14671 standard; cDNA; 3270 BP.
AC AAH14671;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12354.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 12354; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:

```

```

CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 3270 BP; 867 A; 821 C; 852 G; 730 T; 0 U; 0 Other;

```

Alignment Scores:

```

Pred. No.: 4,57e-209 Length: 3270
Score: 2159.00 Matches: 390
Percent Similarity: 89.53% Conservative: 72
Best Local Similarity: 75.58% Mismatches: 54
Query Match: 75.89% Indels: 0
DB: 4 Gaps: 0

```

US-10-029-020-14_COPY_1760_2300 (1-541) x AAH14671 (1-3270)

```

QY 26 HisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsnValThrLeuProIle 45
Db 3 CACGTTCTGGCTGGCCACCGCTAATCCGACGGTTCGCAAAAGAAACATGACTTTGGCTGCGC 62
QY 46 AspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGlnAlaArgGlyGlnVal 65
Db 63 GAGAACGGTCAAAACTTGGTGGATGGAGATCCGAAAGAGAGCAAGCCCAAGGGAAGTC 122
QY 66 ThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeuSerLeuAspPheAsp 85
Db 123 AATGTCITTTGGCCGACAGCTCAGGGTTAATGCGAGAAACCTCTTTTCAGTTGACTTTGAT 182
QY 86 ArgValThrArgThrGluLysIleTyrAspAspHisArgLysPheThrLeuArgGlyLeu 105
Db 183 CGAACACAAAGACAGAAAGATCTATGACACACCCGTAATTTCTACTGAGGATCGCC 242
QY 106 TyrAspGlnAlaGlyArgProSerLeuTrpSerProSerArgLeuAsnGlyValAsn 125
Db 243 TACGACACGCTCTGGGCACCCGACTCTCTGGCTGCCAAGCAGCAGTGTGCGCGTCAAT 302
QY 126 ValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGlyIleMetSerGluArg 145
Db 303 GTACCTATTTCATCCACAGGTCAAAATTCAGCAGCATCCAGCAGGACCACTAGGAGAAA 362
QY 146 MetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAlaAspGlyLysThrTrp 165
Db 363 GTAGATTATGACCGACAGCGGAGGAGATCGTCTCGGGTCTTTGCTGATGGTAAACATGG 422
QY 166 SerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSerGlnArgGlnTyrIle 185
Db 423 AGTTTACATATTAGAAAAAGTCCATGGTTCTTCTGCTTCATAGCCAGCGCAGTACATC 482
QY 186 PheGluPheAspLysAsnAspArgLeuSerSerValThrMetProAsnValAlaArgGln 205
Db 483 TTGATATACGATATGTGGGACCCGCTGTCTGCGCATCACCATGCGCAGTGTGGTGGCCAC 542
QY 206 ThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnProGluGly 225
Db 543 ACCATGACGACCATCGATCCATCTGCTACTACCGCAACATATACAAACCCCGGAAAGC 602

```

[illegible]

Breast cancer; cytostatic; gene therapy; antiseuse therapy; regulated;
drug discovery; clinical medicine; forensic medicine; gene;
chromosome 5q33.3; ds.
Homo sapiens.
WO2003029421-A2.
10-APR-2003.
02-OCT-2002; 2002WO-US031287.
03-OCT-2001; 2001US-0326526P.
14-MAY-2002; 2002US-00144194.
(ORIG-) ORIGENE TECHNOLOGIES INC.
Sun Z, Li X, Fan W, Kovacs KF, Jay G;
WPI: 2003-381623/36.
P-PSDB; ABR58317.
New isolated human differentially-regulated breast cancer polynucleotide
and polypeptide, useful for diagnosing, staging, prognosticating,
preventing and/or treating diseases and conditions relating to breast
cancer.
Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.
The invention relates to isolated polynucleotides which are
differentially-regulated in breast cancer. The methods and compositions
of the present invention are useful for detecting, diagnosing, staging,
monitoring, prognosticating, preventing and/or treating diseases and
conditions relating to breast cancer, and may be used in gene therapy or
antiseuse therapy. They can also be used in research, drug discovery,
clinical medicine and forensic medicine. Sequences given in records
ACC72012-ACC72074 represent polynucleotides of the invention that are
differentially-regulated in breast cancer. NOTE: The sequence data for
this patent did not form part of the printed specification, but was
obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 3_56e-206 | Length: | 9058 |
|------------------------|-----------|---------------|------|
| Score: | 2136.50 | Matches: | 382 |
| Percent Similarity: | 87.06% | Conservative: | 89 |
| Best Local Similarity: | 70.61% | Mismatches: | 69 |
| Query Match: | 75.10% | Indels: | 1 |
| DB: | 7 | Gaps: | 1 |

US-10-029-020-14_COPY_1760_2300 (1-541) x ACC72051 (1-3058)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
||| :
Db 4676 TACCAGCTCTGTATTAATGGTAGTCCTCCTAGGGTGATGTATGCTAAITGGATCGGTATCAGC 4735
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
:
Db 4736 TTCACAGCCAGCCCCCATGCTCCTAGCGGCCACCATCACCCCACCATTTGGACGCTGCAC 4795
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpAtgGlnArylsGluGln 60
:
Db 4796 ATTCCTCGCCTATGGAATGGCTTAACTCCATTGATGGCGGCTCAAGAAGAACAG 4855
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeu 80
:
Db 4856 ATTAAAGCAAAGTCACCATCTTTGGCAGAGACTCCGGTCCATGGAAGAAATCTCTTG 4915
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrrAspAspHisArgLysPhe 100
||| :

4916 TCCATTGACTATGATGAAATATTTCGAGCTGAAAGATCTATGATGACCAACCGGAAGTTC 4975
101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
4976 ACCCTGAGGATCATTTATGACCAAGGTGGCGGCCCTTCCTCTGGCTGCGCAGCGGG 5035
121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
5036 CTGGCAGCTGTCAACGTGTATCTTCTCAATGGCGCTGCTGGCTTTCAGGTGGG 5095
141 IleMetSerGlnArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
5096 GCCATGAGCGAGAGGACAGATCGACCAAGCAAGCGCGCATCTGTGCCGATGTTCTGCT 5155
161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
5156 GACGGGAAAGTGTGGAGCTACTCTCTACCTTGACAGTCCATGTCTCTCTGCTTCAGAGC 5215
181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro 200
5216 CAACGTTCAGTATATTTGAGTATGACTCTCTGACCGCTCTCTGCGCTCACCATGCC 5275
201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
5276 AGCGTGGCCCGGACAGCATGTCACACACACCTCCATCGGTACATCGGTATATTTAC 5335
221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
5336 AACCCGCTGAAAGCATGCTTCGTGTCATCTTGACTACAGTGATGACGGCGCATCTGTG 5395
241 HisThrPheThrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
5396 AAGACTCTCTTTTGGGACCGGACCGAGGTGTTCTACAGTATCGGAATCTTCCAAAG 5455
261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
5456 TTATCAGAGATTGCTTACGACAGTACCGCGCTTCCTTGGGTATGACGAGACCACTGCT 5515
281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
5516 GTCTTGAAGATGGTCAACTCCAAAGTGGGGCTTCTCTGACCATCAGGTACCGGAAG 5575
301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
5576 ATTGGCCCTGGTGGACAGCAGATCTACAGTTCCTCGGAGGAGGATGTCATGCC 5635
321 ArgPheAspTyrAsnTyr--AspAsnSerPheArgValThrSerMetGlnAlaValIle 339
5636 AGGTTTGACTACCTATCATGACAAACAGCTTCGCGCATCGCAAGCATCAAGCCGTCATA 5695
340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGlu 359
5696 AGTCAGACTCCCTCCCGTTGACTCTACCGCTATGATGAGATTTCGGCAAGGTGAA 5755
360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379
5756 CACTTTGGTAAAGTTTGGAGTCATCTATGACATCAACAGATCATCACCACCTGCGGTG 5815
380 MetThrHisThrLysHisPheAspAlaTyrClyArgMetLysGluValGlnTyrGluIle 399
5816 ATGACCTCAGCAACACTTCGACCCATGGCGGATCAAGAGGTCCAGTATGAGATG 5875
400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLys 419
5876 TTCCGGTCCCTCATGTACTGATGACCGGTGCAATATGACATGCGGCGGTGATCAAG 5935
420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
5936 AGGAGTAGTAAACTGGGCGCTATGCAATACCAAGAGTACACCTATGACTACGATGGG 5995
440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAsp 459
5996 GACGGGAGCTCCAGCGGTGGCGGTCAATGACCGCGCCGACCTGGCGCTACAGCTATGAC 6055

QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db 6056 CTTAATGGGAATCTCCACITTAAGAACCCAGGCAACAGTGTGGCTCATGCCCTTGGCG 6115
QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db 6116 TATGACCTCGGGATCGGATACACAGACTCGGGATGTGCAGTACAAAATTGACGAGAT 6175
QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
Db 6176 GGTATCTGTGCCAGAGAGGTCTGACATCTTCAATACAAATTCGAAGGCGCTCTAAACA 6235
QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 6236 AGAGCCTACAACAAGCCAGCGGTGAGTGTCCAGTACCGCTATGATGGCGTAGGACGG 6295
QY 540 Arg 540
Db 6296 CGG 6298

RESULT 11

ACC72052
ID ACC72052 standard; DNA; 9695 BP.

XX ACC72052;

XX 08-JUL-2003 (first entry)

XX BCU0205B gene #SEQ ID 81.

XX Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
XX drug discovery; clinical medicine; forensic medicine; gene;
XX chromosome 5q33.3; ds.
XX Homo sapiens.

XX WO2003029421-A2.
XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031287.
XX 03-OCT-2001; 2001US-0326526P.
XX 14-MAY-2002; 2002US-00144194.

XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX WPI; 2003-381623/36.
XX P-PSDB; ABR58318.

XX New isolated human differentially-regulated breast cancer polynucleotide
XX and polypeptide, useful for diagnosing, staging, prognosticating,
XX preventing and/or treating diseases and conditions relating to breast
XX cancer.

XX Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.

XX The invention relates to isolated polynucleotides which are
XX differentially-regulated in breast cancer. The methods and compositions
XX of the present invention are useful for detecting, diagnosing, staging,
XX monitoring, prognosticating, preventing and/or treating diseases and
XX conditions relating to breast cancer, and may be used in gene therapy or
XX antisense therapy. They can also be used in research, drug discovery,
XX clinical medicine and forensic medicine. Sequences given in records
XX ACC72012-ACC72074 represent polynucleotides of the invention that are
XX differentially-regulated in breast cancer. NOTE: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
 KW anti-lymphatic; antiasthmatic; antileukemia; antileukemia;
 KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
 KW gynaecological; antiinfectivity; immunostimulant; auditory; haemostatic;
 KW gene therapy; FCTR3a; neurexin-like protein; FCTR3f.

OS Homo sapiens.

Key Location/Qualifiers
 5'UTR 1..209
 /*tag= b
 CDS 210..8384
 /*tag= a
 /product= "Human FCTR3F"
 3'UTR 8385..9729
 /*tag= c

WO200166747-A2.

13-SEP-2001.

05-MAR-2001; 2001WO-US007160.

03-MAR-2000; 2000US-0186592P.

03-MAR-2000; 2000US-0186718P.

06-MAR-2000; 2000US-0187293P.

17-MAR-2000; 2000US-0190400P.

07-APR-2000; 2000US-0196018P.

03-JAN-2001; 2001US-0259548P.

(CURA-) CURAGEN CORP.

Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
 Macdougall J, Mishra V, Mezes PS, Rastelli L;

WPI; 2001-596837/67.

P-PSDB; AAU08681.

Novel polypeptides designated as FCTR polypeptides, useful in detection,
 prevention and treatment of a broad range of pathological states.

Claim 9; Page 37-39; 215pp; English.

The invention relates to human FCTR polypeptides, FCTR1-FCTR7, and the
 nucleic acids encoding them. These sequences are useful for the treatment
 or prevention of numerous disorders including myelogenous leukaemia,
 carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
 alloimmune thrombocytopenia, neurological disorders, neurodegenerative
 disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
 -Tooth neuropathy, demyelinating Gardner syndrome, familial
 myelodysplastic syndrome, mental health conditions, familial
 disorders, allergy and infection, bronchial asthma, immunological
 eosinophilia, lung diseases, reproductive disorders, infertility, male
 and female reproductive disorders, deafness, glycoprotein Ia deficiency,
 desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
 disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
 infection, Spino cerebellar ataxia, Plasmodium falciparum parasitaemia,
 Corneal dystrophy-Greenough type I, Corneal dystrophy-lattice type I and
 Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
 the FCTR3a homologue FCTR3f

SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 3,92e-206 | Length: | 9729 |
| Score: | 2136.50 | Matches: | 382 |
| Percent Similarity: | 87.06% | Conservative: | 89 |
| Best Local Similarity: | 70.61% | Mismatches: | 69 |
| Query Match: | 75.10% | Indels: | 1 |
| DB: | 5 | Gaps: | 1 |

US-10-029-020-14_COPY_1760_2300 (1-541) x AAS14089 (1-9729)

| | | | |
|----|------|--|------|
| QY | 1 | TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuLaAsnGlyMetGluValala | 20 |
| DB | 5361 | TACCAGCTCTGTAATAATGATGCTGAGGGTGAATGATGCTAATGGATGGGTATCAGC | 5420 |
| QY | 21 | LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn | 40 |
| DB | 5421 | TTCCACAGCGAGCCCATGCTTAGCGGCGACCATCACCCACCATTGGACGCTGCAAC | 5480 |
| QY | 41 | ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln | 60 |
| DB | 5481 | ATCTCCCTGCTATGAGGAATGCTTAAACTCCATTGAGTGGCGCTTAAGAAAGAAACAG | 5540 |
| QY | 61 | AlaArgGlyGlnValThrValPheGlyArgArgValHisAsnArgAsnLeuLeu | 80 |
| DB | 5541 | ATTAAAGGCAAGTCACCATCTTTGCGAGGAAGTCCGGGTCCATGGAAGAAATCTCTTG | 5600 |
| QY | 81 | SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe | 100 |
| DB | 5601 | TCCATTGACTATGATCGAAATATTCGGAATGAAAGATCTATGATGACCCAGGAGTTC | 5660 |
| QY | 101 | ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg | 120 |
| DB | 5661 | ACCTGAGGATCATTTATGACGAGTGGCGCCCTTCTCTGGTGCACCGAGGCGG | 5720 |
| QY | 121 | LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly | 140 |
| DB | 5721 | CTGGCAGCTGTCAACGTGTCTACTTCTTCATGATGGCGCTGGCTGGGCTTCAGCGTGG | 5780 |
| QY | 141 | IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla | 160 |
| DB | 5781 | GCCATGAGCGAGGAGACACATCGAACAAGCCGCGCATCGTCCCGCATGTTCGT | 5840 |
| QY | 161 | AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer | 180 |
| DB | 5841 | GACGGAAAGTGTGGAGCTACTTCTTACCTTGAAGTCCATGCTCTCTCTCCAGAGC | 5900 |
| QY | 181 | GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro | 200 |
| DB | 5901 | CAAGTCAATATATATTTGATGATGATCTCTCTGACCGCTCTTGGCGTACCAGTCC | 5960 |
| QY | 201 | AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr | 220 |
| DB | 5961 | AGCGTGGCGCGCACAGCATGTCCACACACACCTCCATCGGTACATCCGTAATATTAC | 6020 |
| QY | 221 | GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu | 240 |
| DB | 6021 | AACCGCTGAAAGCAATGCTTCGGTCTATCTTGAATGATGATGATGATGATGATGAT | 6080 |
| QY | 241 | HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys | 260 |
| DB | 6081 | AAGACCTCTTTTGGGCGAGCGAGCGAGTGTCTTCAAGTATGGAAACTCTCCAG | 6140 |
| QY | 261 | LeuAlaGluThrLeuTyrAspThrLysValSerPheThrTyrAspGluThrAlaGly | 280 |
| DB | 6141 | TTATCAGAGATGTGCTACGACGATGACCGCGTCACTTCGGGTATGACGACCACTGGT | 6200 |
| QY | 281 | MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln | 300 |
| DB | 6201 | GTCTTGAAGATGTCACCTCCAAAGTGGGGGCTTCTCTGACCATCAGTACCGGAAG | 6260 |
| QY | 301 | IleGlyProLeuLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla | 320 |
| DB | 6261 | ATTGGCCCCCTGGTGGACAGCATCTACAGGTTCTCCGAGGAGGATGCTGATGCC | 6320 |
| QY | 321 | ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle | 339 |
| DB | 6321 | AGGTTGACTACCTATCATGACACAGCTTCGGCATCGCAAGATCAAGCCGTCATA | 6380 |
| QY | 340 | AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu | 359 |
| DB | 6381 | AGTGAGACTCCCTCCCGTTGACCTCTACCGCTATGATGAGATTTCTGGCAAGTGGAA | 6440 |

QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIlelleThrThrAlaVal 379
Db 6441 CACTTTGGTAAGTTGGAGTCATCTATTATGACATCAACCATCATCAACACTCCCGTG 6500
QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db 6501 ATGACCTCAGCAACACATTCACACCCATGGCGGATCAAGGAGTCCAGTATGAGATG 6560
QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVallys 419
Db 6561 TTCCGGTCCCTCATGTACTGATGACGGTGCATATGACAGCATGGCGAGGTGATCAAG 6620
QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db 6621 AGGGAGCTAAACTGGGGCCCTATGCCAATACACCAAGATACACCTATGACATGAC 6680
QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAsp 459
Db 6681 GACGGCAGCTCCAGAGCGTGGCGTCAATGACCGCCGACCTGGCGCTACAGCTATGAC 6740
QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db 6741 CTTAATGGGAATCTCCACTTACTTAACCCAGGCAACAGTGTGGCCTCATGCGCTTGC 6800
QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db 6801 TATGACCTCCGGATCGGATACACACTCGGGAATGTGAGTACAAATGACGAGAT 6860
QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
Db 6861 GCGTATCTGTCCAGAGAGGCTGTGACATCTCGAATACAAATCCAAAGGCGCTCTCA 6920
QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 6921 AGAGCTTACAAACAAAGGCGGAGTGGAGTGTCCAGTACCGCTATGATGGGTAGGAC 6980
QY 540 Arg 540
Db 6981 CGG 6983
RESULT 13
ID ADB32028 standard; cDNA; 9729 BP.
XX AC ADB32028;
XX DT 04-DEC-2003 (first entry)
XX DE Human FCTR3f cDNA.
XX KW Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW lattice corneal dystrophy.
XX OS Homo sapiens.
XX PN US2003087816-A1.
XX PD 08-MAY-2003.
XX PF 05-MAR-2001; 2001US-00800198.
XX

03-MAR-2000; 2000US-0186592P.
(VERM/) VERMET C.
PA (FERN/) FERNANDES E.
PA (SHIM/) SHIMKETS R.
PA (HERR/) HERMANN J.
PA (MAJU/) MAJUMDER K.
PA (MACD/) MACDOUGALL J.
PA (MISH/) MISHRA V.
PA (MEZE/) MEZES P S.
PA (RASP/) RASTELLI L.
XX
PI Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI: 2003-625633/59.
DR P-PSDB; ADB32029.
XX
PT New FCTRX polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTRX-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.
XX
PS Claim 9; Page 34-37; 155pp; English.
XX
CC The invention relates to FCTRX polypeptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTRX polypeptide, such as colorectal cancer,
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, nerve trauma, familial myelodysplastic
CC neurodegenerative disorders, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
CC conditions, immunological disorders, allergy and infection, asthma, lung
CC diseases, male and female reproductive disorders, deafness, glycoprotein
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, plasmodium falciparum infection,
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTRX polypeptide of the invention.
XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 3.92e-206 Length: 9729
Score: 2136.50 Matches: 382
Percent Similarity: 87.06% Conservative: 89
Best Local Similarity: 70.61% Mismatches: 69
Query Match: 75.10% Indels: 1
DB: Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x ADB32028 (1-9729)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 5361 TACCAGCTCTCTAATAATGTTACCTGAGGTGATGTATGTTAATGGATGGTATCAGC 5420
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5421 TTCACAGCGAGCCCACTGCTTACCGGCGACCATCCCCACCATTGGAGCTGCAAC 5480
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5481 ATCTCCCTGCCTATGAGAAATGGCTTAACTCCATTGATGGCGCTTAAAGGAACAG 5540
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5541 ATTAAGGCAAGTACCACATCTTTGGCAGGAAGTCCCGGTTCATGGAGAAATCTCTTG 5600
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100

5601 Db TCCATTGACTATGATCGGAAATATTCCGGACTGAAAGATCTATGATGATGACCAACCGGAAGTTC 5660
 QY ThrLeuArgIleLeuTyrAspGlnaIaGlyArgProSerLeuTyrPsrProSerSerArg 120
 Db ACCCTGAGATCATTTATGACAGAGTGGCGCCCTTCCTCTGGCTGCCAGCAGCGGG 5720
 QY LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
 Db CTGGCAGCTGTCAACGTGTCTACTCTTCAATGGCGCTGGCTGGCTTACGCTGGG 5780
 QY IleMetSerGluArgMetGluTyrAspGlnaIaGlyArgIleThrSerArgIlePheAla 160
 Db GCCATGACGAGACAGACATCGACAAGACGCGCATCGTGTCCCGCATGTCCTGC 5840
 QY AspGlyLysThrTrpSerTyrThrTyrLeuGlnLysSerMetValLeuLeuHisSer 180
 Db GACGGAAAGTGTGGAGTACTCTCTGCTGCAAGTCCATGGTTCCTCTGCTTCAGAGC 5900
 QY GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
 Db CAACGTCAATATATTGATATGATCTCTCTGACCGCTCTCTGCGGTCAACATGCC 5960
 QY AsnValIaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
 Db AGCTGGCCCGGACACAGATGTCACACACACCTCCATCGCTCATCGCTTAATTTAC 6020
 QY GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
 Db AACCGCTGAAAGCAATGCTTCGTGATCTTTGACTACAGTATGATGACGCGCCGCTCG 6080
 QY HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
 Db AAGACCTCTTTTGGGCACCGGACCGAGTGTCTTACAAGTATGGAACATCTCCAG 6140
 QY LeuAlaGlnThrLeuTyrAspThrThrLysValSerPheThrTyrTyrAspGluThrAlaGly 280
 Db TTATCAGAGATGTCTACGACAGTACCGCGCTCACCTTCGGGTATGACGAGACCATG 6200
 QY MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
 Db GTCTTGAGATGTCACCTCCAAAGTGGGGGCTTCTCTGACCATCAGGTACCGGAAG 6260
 QY IleGlyProLeuLeuAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
 Db ATTGGCCCCCTGTGGACAGACAGATCTACAGGTCTCCGAGGAGGCAATGTCATGCC 6320
 QY ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339
 Db AGTGGACATCCCTCCCGTGTGACCTTCACCGCTATGATGATGATTTCTGGCAAGTGA 6380
 QY AsnGluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGlu 359
 Db AGTGAGACTCCCTCCCGTGTGACCTTCACCGCTATGATGATGATTTCTGGCAAGTGA 6440
 QY GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379
 Db CACTTTGGTAAGTTTGGAGTCACTATTATGATCATCAACAGATCATCACCATGCCGTG 6500
 QY MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluLe 399
 Db ATGACCTCAGCAAAACACTTCGACACCCATGCGCGGATCAAGAGAGTCCAGTATGAG 6560
 QY PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLys 419
 Db TTCCGCTCCTCATGTACTGGATGACGGTGCATATGACAGTATGGGAGGGGTGATCAAG 6620
 QY LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
 Db AGGAGACTAAACCTGGGGCCCTATGCCAATACCAAGAGTACACCTATGACTACGATGG 6680
 QY AspGlyGlnLeuGlnThrValSerLeuAsnAspLysProLeuTyrArgTyrSerTyrAsp 459

| | | | |
|-----------|--------------------------------|---|------|
| Db | 6681 | GACGGCAGCTCCAGAGCGTGGCCGTCAATGACCGCCGACCTGGCCCTACAGCTATGAC | 6740 |
| Qy | 460 | LeuAnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg | 479 |
| Db | 6741 | CTTAATGGGAATCTCCACTTACTGAACCCAGGCAACAGTGTGGCGCTCATGTGCCCTTGGCG | 6800 |
| Qy | 480 | TyrAspIleArgAspArgGlieThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp | 499 |
| Db | 6801 | TATGACCTCCGGATCGGATACAGACTCGGGATGTGCAGTACAAATTCGACGAT | 6860 |
| Qy | 500 | GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle | 519 |
| Db | 6861 | GGCTATCTGTGCCAGAGAGGGTCTGACATCTTCGATACAAATTCCAAGGCGCTCCCTAAC | 6920 |
| Qy | 520 | LysAlaTyrAsnArgAlaGlySerTTrpSerValArgTyrArgTyrAspGlyLeuGlyArg | 539 |
| Db | 6921 | AGAGCTACAAACAGGCGCGGTGGAGTGTCCAGTACCGCTATGATGCGGTAGACGG | 6980 |
| Qy | 540 | Arg 540 | |
| Db | 6981 | CGG 6983 | |
| RESULT 14 | | | |
| AAS14085 | | | |
| ID | AAS14085 | standard; DNA; 9826 BP. | |
| XX | AC | AAS14085; | |
| XX | DT | 18-DEC-2001 (first entry) | |
| XX | DE | Human FCTR3b DNA sequence. | |
| KW | KW | Human; FCTRX; myelogenous leukaemia; carcinoma; melanoma; glioma; ds; | |
| KW | KW | ascrocytoma; congenital neonatal alloimmune thrombocytopaenia; infection; | |
| KW | KW | neurological disorder; neurodegenerative disorders; nerve trauma; | |
| KW | KW | familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy; | |
| KW | KW | demyelinating Gardner syndrome; familial myelodysplastic syndrome; | |
| KW | KW | mental health condition; immunological disorder; allergy; infertility; | |
| KW | KW | bronchial asthma; Avellino type eosinophilia; lung disease; deafness; | |
| KW | KW | reproductive disorder; reproductive disorder; glycoprotein Ia deficiency; | |
| KW | KW | desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide; | |
| KW | KW | gastric disorders; pancreatic disease; Schistosoma mansoni infection; | |
| KW | KW | Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes; | |
| KW | KW | Corneal dystrophy-Greenow type I; Corneal dystrophy-lattice type I; | |
| KW | KW | Reis-Bucklers corneal dystrophy type I; Corneal dystrophy-lattice type I; | |
| KW | KW | antiallergic; antiasthmatic; antiinfertility; immunosuppressive; | |
| KW | KW | antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological; | |
| KW | KW | gynaecological; antiinfertility; immunostimulant; auditory; haemostatic; | |
| XX | XX | gene therapy; FCTR3b; neurestin-like protein. | |
| OS | OS | Homo sapiens. | |
| XX | XX | | |
| Key | | Location/Qualifiers | |
| 5'UTR | 1..279 | | |
| FT | FT | /*tag= b | |
| FT | FT | 280..8481 | |
| CDS | | /*tag= a | |
| FT | FT | /product= "Human FCTR3b" | |
| FT | FT | 8482..9826 | |
| FT | FT | /*tag= c | |
| XX | XX | | |
| XX | XX | | |
| PN | WO200166747-A2. | | |
| XX | XX | | |
| PD | 13-SEP-2001. | | |
| XX | XX | | |
| PF | 05-MAR-2001; 2001WO-US0007160. | | |
| XX | XX | | |
| PR | 03-MAR-2000; 2000US-0186592P. | | |
| PR | 03-MAR-2000; 2000US-0186718P. | | |
| PR | 06-MAR-2000; 2000US-0187293P. | | |
| PR | 06-MAR-2000; 2000US-0187294P. | | |
| PR | 17-MAR-2000; 2000US-0190400P. | | |
| PR | 07-APR-2000; 2000US-0196018P. | | |

| | | | | | |
|---|---|------|------|---|------|
| PR | 03-JAN-2001; 2001US-0259548P. | QY | 141 | IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla | 160 |
| XX | | Db | 5878 | GCCATGAGGAGGAGGAGACATCGACAAGACCGCATCGTCTCCGCGCATGTTCGT | 5937 |
| PA | (CURA-) CURAGEN CORP. | | | | |
| XX | Vernet CM, Fernandes E, Shinkets RA, Herrmann JU, Majumder K; | QY | 161 | AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer | 180 |
| PI | Maddougall J, Mishra V, Mezes PS, Rastelli L; | Db | 5938 | GACGGGAAAGTGTGGAGCTACTCTACCTTGGCAAGTCCATGGTCTCTCTCTTCAGAC | 5997 |
| XX | WPI; 2001-596837/67. | QY | 181 | GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerValThrMetPro | 200 |
| DR | P-PSDB; AAU08680. | Db | 5998 | CAACGTCAGTATATATTTGAGTATGACTCTCTGACCGCTCTCTGCGCTACCATGCC | 6057 |
| XX | | QY | 201 | AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr | 220 |
| DR | | Db | 6058 | AGCGTGGCGGCACACGATGTCACACACACCTCCATCGGTACATCGGTATATTTAC | 6117 |
| XX | Claim 9; Page 33-35; 215pp; English. | QY | 221 | GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu | 240 |
| XX | The invention relates to human FCRLX polypeptides, FCRLX-FCRL7, and the | QY | 6118 | AACCCGCTGAAAGCAATGCTTCGGTCACTTTTGAATACAGTATGACGGCCGCTCTG | 6177 |
| CC | nucleic acids encoding them. These sequences are useful for the treatment | Db | 241 | HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrIlystYrGlyLysLeuSerLys | 260 |
| CC | or prevention of numerous disorders including myelogenous leukaemia, | QY | 6178 | AGACCTCTCTTTTGGCAGCCGACCCAGGTGTCTTACAAGTATGGAAACTCTCCAG | 6237 |
| CC | carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal | Db | 261 | LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly | 280 |
| CC | allimmune thrombocytopaenia, neurological disorders, neurodegenerative | QY | 6238 | TTATCAGAGATTGCTACGACAGTACCGCGTCACTTCGGGTATGACGAGACCATGGT | 6297 |
| CC | disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie | Db | 281 | MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln | 300 |
| CC | -Tooth neuropathy, demyelinating Gardner syndrome, familial | QY | 6298 | GTCTTGAAGATGTCACCTCCAAAGTGGGGGCTTCTCCGTCACCATCAGTACCGAAG | 6357 |
| CC | myelodysplastic syndrome, mental health conditions, immunological | Db | 301 | IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla | 320 |
| CC | disorders, allergy and infection, bronchial asthma, Avellino type | QY | 6358 | ATTGGCCCCCTGGTGGACAGAGATCTACAGGTCTTCCGAGGAAGCATGGTCAATGCC | 6417 |
| CC | sosinophilia, lung diseases, reproductive disorders, infertility, male | QY | 321 | ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle | 339 |
| CC | and female reproductive disorders, deafness, glycoprotein Ia deficiency, | Db | 6418 | AGGTTTGACTACACCTATCATGACACAGCTTCGACATCAACAGATCATCACCCTCGCTG | 6477 |
| CC | desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric | QY | 340 | AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu | 359 |
| CC | disorders, pancreatic diseases such as diabetes, Schistosoma mansoni | Db | 6478 | AGTGAGACTCCCTCCCGCTTGACCTCTACCGTATGATGAGATTTCTGGCAAGGTGAA | 6537 |
| CC | infection Spino cerebellar ataxia, plasmodium falciparum parasitaemia, | QY | 360 | GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal | 379 |
| CC | Corneal dystrophy-Greenow type I, Corneal dystrophy-lattice type I and | Db | 6538 | CACCTTGGTAAGTTTGGAGTCATCTATTATGACATCAACAGATCATCACCCTCGCTG | 6597 |
| CC | Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding | QY | 380 | MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle | 399 |
| CC | FCRL3b, a neurastin-like protein | Db | 6598 | ATGACCTCAGCAAAACATTCGACACCCATGGGCGGATCAAGGAGGTCCAGTATGAGATG | 6657 |
| CC | Sequence 9826 BP; 2567 A; 2643 C; 2479 G; 2137 T; 0 U; 0 Other; | QY | 400 | PheArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValLys | 419 |
| XX | | Db | 6658 | TTCCGGTCCCTCATGTACTGATGACGGTGAATATGACACATGCGGCGGCTGATCAAG | 6717 |
| Alignment Scores: | | QY | 420 | LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla | 439 |
| Pred. No.: | 3 97e-206 | Db | 6718 | AGGAGAGCTAAAACCTGGGCGCTTATGCCAATACCAAGAGTACACCTATGATGATGGG | 6777 |
| Score: | 2136.50 | QY | 440 | AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAsp | 459 |
| Percent Similarity: | 87.06% | Db | 6778 | GACGGGAGCTCCAGAGGTGGCGTCAATGACCGCCGCTCACTGGCGGTACAGTATGAC | 6837 |
| Best Local Similarity: | 70.61% | QY | 460 | LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg | 479 |
| Mismatch: | 69 | Db | 6838 | CTTATGGGATCTCCACTTACTGAACCCAGCAACAGTGTGGCGCTCATGCCCTTGGCG | 6897 |
| Indels: | 1 | QY | 480 | TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp | 499 |
| Gaps: | 5 | Db | 6898 | TATGACCTCCGGATCGGATTAACAGACTCGGGGATGTCAGTACAAATATGACCATG | 6957 |
| Query Match: | 5 | QY | 500 | GlyPheLeuArgGlnArgGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle | 519 |
| DB: | | Db | | | |
| US-10-029-020-14_COPY_1760_2300 (1-541) x AAS14085 (1-9826) | | | | | |
| QY | 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla | 20 | | | |
| Db | 5458 TACCAGCTCTGTAATAATGTTACCTGAGGTGATGATGCTAATGGATGGGTATCAGC | 5517 | | | |
| QY | 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn | 40 | | | |
| Db | 5518 TTCACAGCGAGCCCGATGCTCCTAGCGGCGACCAATCACCCACCATGTCGCGTGCAC | 5577 | | | |
| QY | 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrPArgGlnArgLysGluGln | 60 | | | |
| Db | 5578 ATCTCCCTGCTATGGAGATGGCTTAACTCCATGAGTGGCGCTAAGAAAGAAACAG | 5637 | | | |
| QY | 61 AlaArgGlyGlnValThrValPheGlyArgLeuArgValHisAsnArgAsnLeuLeu | 80 | | | |
| Db | 5638 ATTAAGGCAAGTCAACATCTTTGGCAGGAGCTCCGGGTCCATGGGAGAAATCTCTTG | 5697 | | | |
| QY | 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe | 100 | | | |
| Db | 5698 TCCATTTGATGATCGAAATATTCGGACTGAAAGATCTATGATGACCCGGAAGTTC | 5757 | | | |
| QY | 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrProSerProSerArg | 120 | | | |
| Db | 5758 ACCCTGAGATCATTTATGACAGGTGGCGGCCCTTCTCTGCTGCCAGCAGCGGG | 5817 | | | |
| QY | 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly | 140 | | | |
| Db | 5818 CTGCGAGCTGTCAAGTGTCTTCTTCAATGGCGCTCTGCTGGCTTCAGGCTGGG | 5877 | | | |

Mon Aug 16 09:01:10 2004

```

Db      6118 AACCGCGCTGAAGCAATGCTTCGGTCATCTTTGACTACAGTGATGACGGCGCATCTCG 6177
Qy      241  HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db      6178 AAGACCTCTCTTTTGGGACCGACCGCGAGGTGTTCTACAGTATGGGAAACTCTCCAAG 6237
Qy      261  LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db      6238 TTATCAGAGATTGTCTACGACAGTACCGCGCTCACCTTCGGGTATGACGAGACCCTGGT 6297
Qy      281  MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db      6298 GTCTTGAAGATGGTCAACCTCAAAAGTGGGGCTTCTCTGCACCATCAGGTACCGGAAG 6357
Qy      301  IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db      6358 ATGGCCCCCTGGTGACAGAGATCTACAGGTTCTCCGAGGAGGATGGTCAATGCC 6417
Qy      321  ArgPheAspTyrAsnTyr--AspAsnSerPheArgValThrSerMetGlnAlaValIle 339
Db      6418 AGGTTTGACTATACCTATCATGACAAACAGCTCCCGATCGCAAGCATCAAGCCCGTCATA 6477
Qy      340  AsnGluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGlu 359
Db      6478 AGTGAGACTCCCTCCCGTTGACCTCTACCGCTATGATGAGATTCTGGCAAGGTGAA 6537
Qy      360  GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379
Db      6538 CACTTTGGTAAGTTGGAGTCACTATTATGACATCAACACAGATCATCACCTACCTGCGGTG 6597
Qy      380  MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db      6598 ATGACCCCTCAGCAAAACACTTCGACACCCATGGGGGATCAAGGAGTCCAGTATGAGATG 6657
Qy      400  PheArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValValLys 419
Db      6658 TTCCGGTCCCTCATGTACTGTGATGACCGTGCAATATGACGATGGGCGGTGATCAAG 6717
Qy      420  LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db      6718 AGGAGCTAAACTGGGGCCCTATGCCAATACCACGAAGTACACCTATGACTACGATGGG 6777
Qy      440  AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAsp 459
Db      6778 GACGGGCGAGCTCCAGAGCGTGGCGGTCAATACCCCGCCGACCTGGCGGTACAGCTATGAC 6837
Qy      460  LeuAsnGlyAsnLeuHisLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db      6838 CTTAATGGGAATCTCCACTTACTGAACCCAGGCAACAGTGTGGCGCTCATGCCCTTGGCG 6897
Qy      480  TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db      6898 TATGACCTCCGGATCGGATAACACAGACTCGGGGATGTGCAGTACAAAATTGACAGCAT 6957
Qy      500  GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
Db      6958 GGCTATCTGTGCCAGAGAGGGTCTGACATCTTCGATATACATTCACAGGGCTCTCAACA 7017
Qy      520  LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539
Db      7018 AGAGCTTACACAAAGCCAGCGGGTGGAGTGTCCAGTACCGCTATGATGGCGTAGGACGG 7077
Qy      540  Arg 540
Db      7078 CGG 7080

```

Search completed: August 14, 2004, 02:33:50
Job time : 1049.68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 523.097 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_1490_1750
Perfect score: 1384
Sequence: 1 KINRIQVTSGEISVAGA.....SSKDDVTITNLSASGAFY 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTC_spool/US10029020/runat_06082004_112218_29331/app query.fasta_l.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTEWT=pto -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020@cgn1.1.2156@runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*

| | |
|-----|--|
| 1: | /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.* |
| 2: | /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.* |
| 3: | /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.* |
| 4: | /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.* |
| 5: | /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.* |
| 6: | /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.* |
| 7: | /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.* |
| 8: | /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.* |
| 9: | /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.* |
| 10: | /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.* |
| 11: | /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.* |
| 12: | /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.* |
| 13: | /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.* |
| 14: | /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.* |
| 15: | /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.* |
| 16: | /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.* |
| 17: | /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.* |
| 18: | /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.* |
| 19: | /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
|------------|-------|-------------|--------|----|-------------|

ALIGNMENTS

RESULT 1

US-10-383-201-43

; Sequence 43, Application US/10383201

; Publication No. US20040029226A1

; GENERAL INFORMATION:

; APPLICANT: Alsobrook II, John et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS

; FILE REFERENCE: 21402-568A

; CURRENT APPLICATION NUMBER: US/10/383,201

; CURRENT FILING DATE: 2003-03-06

; PRIOR FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 10/029020

; PRIOR FILING DATE: 2002-03-20

; PRIOR APPLICATION NUMBER: 60/372,022

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/389,143

; PRIOR FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 60/391,779

; PRIOR FILING DATE: 2002-06-26

; PRIOR APPLICATION NUMBER: 60/410,755

; PRIOR FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: 60/412,957

| | | | | | | |
|----|-------|-------|---------|----|----------------------|-----------------------|
| 1 | 1384 | 100.0 | 8354 | 13 | US-10-383-201-43 | Sequence 43, Appl |
| 2 | 1384 | 100.0 | 8354 | 13 | US-10-029-020-13 | Sequence 13, Appl |
| 3 | 1384 | 100.0 | 8438 | 13 | US-10-042-865-1 | Sequence 1, Appl |
| 4 | 1379 | 99.6 | 8355 | 13 | US-10-383-201-55 | Sequence 55, Appl |
| 5 | 935.5 | 67.6 | 8473 | 17 | US-10-038-854-39 | Sequence 39, Appl |
| 6 | 935.5 | 67.6 | 8487 | 17 | US-10-038-854-41 | Sequence 41, Appl |
| 7 | 935.5 | 67.6 | 8645 | 17 | US-10-038-854-37 | Sequence 37, Appl |
| 8 | 935.5 | 67.6 | 8675 | 17 | US-10-038-854-35 | Sequence 35, Appl |
| 9 | 898 | 64.9 | 12880 | 16 | US-10-295-027-927 | Sequence 927, Appl |
| 10 | 865.5 | 62.5 | 8689 | 9 | US-09-808-602-78 | Sequence 78, Appl |
| 11 | 865.5 | 62.5 | 8689 | 10 | US-09-800-198-66 | Sequence 66, Appl |
| 12 | 864.5 | 62.5 | 8797 | 9 | US-09-808-602-74 | Sequence 74, Appl |
| 13 | 864.5 | 62.5 | 8797 | 9 | US-09-808-602-77 | Sequence 77, Appl |
| 14 | 864.5 | 62.5 | 8797 | 10 | US-09-800-198-62 | Sequence 62, Appl |
| 15 | 864.5 | 62.5 | 8797 | 10 | US-09-800-198-65 | Sequence 65, Appl |
| 16 | 863.5 | 62.4 | 6560 | 9 | US-09-808-602-76 | Sequence 76, Appl |
| 17 | 863.5 | 62.4 | 6560 | 10 | US-09-800-198-64 | Sequence 64, Appl |
| 18 | 863.5 | 62.4 | 9058 | 16 | US-10-144-194A-79 | Sequence 79, Appl |
| 19 | 863.5 | 62.4 | 9695 | 16 | US-10-144-194A-81 | Sequence 81, Appl |
| 20 | 863.5 | 62.4 | 9729 | 9 | US-09-808-602-12 | Sequence 12, Appl |
| 21 | 863.5 | 62.4 | 9729 | 10 | US-09-800-198-12 | Sequence 12, Appl |
| 22 | 863.5 | 62.4 | 9826 | 9 | US-09-808-602-7 | Sequence 7, Appl |
| 23 | 863.5 | 62.4 | 9826 | 10 | US-09-800-198-7 | Sequence 7, Appl |
| 24 | 862.5 | 62.3 | 8409 | 9 | US-09-808-602-79 | Sequence 79, Appl |
| 25 | 862.5 | 62.3 | 8409 | 10 | US-09-800-198-67 | Sequence 67, Appl |
| 26 | 861.5 | 62.2 | 8575 | 13 | US-10-072-012-143 | Sequence 143, Appl |
| 27 | 389 | 28.1 | 3310 | 13 | US-09-823-245A-482 | Sequence 482, Appl |
| 28 | 270.5 | 19.5 | 3614 | 13 | US-10-342-887-1743 | Sequence 1743, Appl |
| 29 | 270.5 | 19.5 | 3614 | 13 | US-10-172-118-1743 | Sequence 1743, Appl |
| 30 | 160 | 11.6 | 450 | 15 | US-10-029-386-16526 | Sequence 16526, Appl |
| 31 | 160 | 11.6 | 555 | 15 | US-10-029-386-2826 | Sequence 2826, Appl |
| 32 | 159 | 11.5 | 457 | 13 | US-10-027-632-183606 | Sequence 183606, Appl |
| 33 | 159 | 11.5 | 457 | 16 | US-10-027-632-183606 | Sequence 183606, Appl |
| 34 | 133.5 | 9.6 | 182 | 15 | US-10-029-386-16081 | Sequence 16081, Appl |
| 35 | 133.5 | 9.6 | 519 | 15 | US-10-029-386-2381 | Sequence 2381, Appl |
| 36 | 131.5 | 9.5 | 3750 | 15 | US-10-156-761-22 | Sequence 22, Appl |
| 37 | 131.5 | 9.5 | 9025608 | 15 | US-10-156-761-1 | Sequence 1, Appl |
| 38 | 119.5 | 8.6 | 4716 | 13 | US-10-282-122A-33231 | Sequence 33231, Appl |
| 39 | 117.5 | 8.5 | 3957 | 16 | US-10-369-493-34930 | Sequence 34930, Appl |
| 40 | 115 | 8.3 | 31960 | 15 | US-10-114-170-11 | Sequence 11, Appl |
| 41 | 111.5 | 8.1 | 4233 | 16 | US-10-398-221-1343 | Sequence 1343, Appl |
| 42 | 111.5 | 8.1 | 4535 | 16 | US-10-418-861B-54 | Sequence 54, Appl |
| 43 | 110.5 | 8.0 | 5394 | 16 | US-10-398-221-3837 | Sequence 3837, Appl |
| 44 | 109.5 | 7.9 | 6392 | 16 | US-10-398-221-689 | Sequence 689, Appl |
| 45 | 109.5 | 7.9 | 6504 | 16 | US-10-398-221-2219 | Sequence 2219, Appl |

Mon Aug 16 09:01:09 2004

```

; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 1,99e-170 Length: 8354
Score: 1384.00 Matches: 261
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-383-201-43 (1-8354)

QY 1 LysileAsnArgileArgGlnValThrThrSerGlyGluileSerLeuValAlaGlyAla 20
Db 4502 AAGATCAACCGCATCAGGAGGTCACCATAGTGAGAGATCTCATCTGCTGGGGCC 4561

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4562 CCCAGTGGCTGTGACTGATAAATGATGCAACTGTGATTTCTTCTGAGACGATGGT 4621

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4682 CTCTACGTGGCGACCTTGGGAACATCGGAATTCGGTTTATCCGGAAGAACCAAGCCTTTC 4741

QY 81 LeuAsnThrGlnAsnMetTyrrGluLeuSerSerProilleAspGlnGluLeuTyrrPhe 100
Db 4742 CTCAACACCCAGAACATGATGAGCTGTCTTACCAATTGACCGAGGAGCTCTATCTGTTT 4801

QY 101 AspThrThrGlyLysHisLeuTyrrThrGlnSerLeuProThrGlyAspTyrrLeuTyrrAsn 120
Db 4802 GATACCAACCGGCAAGCAGCCTGTACCCCAAGCCCTGCCACAGAGACTACTCTGACAAAC 4861

QY 121 PheThrTyrrThrGlyAspGlyAspPheThrLeulleThrAspAsnAsnGlyAsnMetVal 140
Db 4862 TTCACCTACCTACCTGGGCGGCGGACATCATCATCATCAGACACAAATGGCAACATGGTA 4921

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrrLeuValProAspGlyGlnVal 160
Db 4922 AATGTCCCGGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGTG 4981

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrGlnGlyHisGlu 180
Db 4982 TACTGGGTGACCATGGGCACCAACAGTGCACATCAAGAGTGTGACCAACAGGACACGAG 5041

QY 181 LeuAlaMetMetTyrrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5042 TTGGCCATGATGACATACCATGGCAATTCGGCCCTCTGGCAACCAAAAGCAATGAAAC 5101

QY 201 GlyTrpThrThrPheTyrrGluTyrrAspSerPheGlyArgGluLeuThrAsnValThrPhePro 220
Db 5102 GGATGGACAAACATTTTATGAGTACGACGCTTTGGCCGCTGACAAATGTGACCTTCCT 5161

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240

```

```
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4742 CTCACACCCAGAACATGATGAGCTGCTTCACCAATTGACCGAGGAGCTCTATCTGTTT 4801
QY 101 AspThrThrGlyHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4802 GATACACCCGGCAAGCACTGTACACCCAAAGCTGCCCAAGGAGACTACCTGTACAAC 4861
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4862 TTCACCTACACTGGGAGCGGACATCACACTATCACAGACAAATGGCAACATGGTA 4921
QY 141 AsnValArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
Db 4922 AATGTCCGCGAGACTCTACTGGATGCCCTCTGGCTGGTGGTCCAGATGCCAGGTG 4981
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuIleSerValThrGlnGlyHisGlu 180
Db 4982 TACTGGTGACCATGGGACCAACAGTGCACTCAAGAGTGTGACCACAAAGGACAGAG 5041
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5042 TTGGCCATGATGACATACCATGCAATTCGGGCTTCTGGCAACCAAGCAATGAAGAA 5101
QY 201 GlyTrpThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5102 GGATGGACAACATTTTATGAGTACGACAGCTTTGGCGCGCTGACAAATGTGACCTTCC 5161
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5162 ACTGGCCAGGTGAGCAGTTTCGAAAGTGATACAGACAGTTCAGTGCATGTCAGGTAGAG 5221
QY 241 ThrSerSerLysAspAspValThrIleThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 5222 ACCTCCAGCAGATGATGTACCATACCAACCAACCTGTGCTCGCTTCAGGCGCTTCTAC 5281
QY 261 Thr 261
Db 5282 ACA 5284

RESULT 3
US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldos, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
```

```
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1

Alignment Scores:
Pred. No.: 2,02e-170 Length: 8438
Score: 1384.00 Matches: 261
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
```

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-042-865-1 (1-8438)

```
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4508 AAGATCAACCGCATCAGGAGGTCCACCATGTAGGAGAGATCTCACTGTTGCTGGGCC 4567
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4568 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTCTGGAGACCATG 4627
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4628 TATGCCAAGAGTGAAGAGTTAAATACCCATCTTCCTTGGCTGTGTGCTGATGGGAG 4687
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4688 CTCCTACGTGGCGGACCTTGGGAAACATCCGAATTCGGTTTATCCGAGAGCAACAGC 4747
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4748 CTCACACCCAGAACATGATGAGCTGCTTCACCAATTGACAGGAGCTCTATCTGTTT 4807
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4808 GATACACCCGGCAAGCACTGTACACCCAAAGCTGCCCAAGGAGACTACCTGTACAAC 4867
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4868 TTCACCTACACTGGGAGCGGACATCACACTATCACAGACAAATGGCAACATGGTA 4927
QY 141 AsnValArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
Db 4928 AATGTCCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCGGAG 4987
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuIleSerValThrGlnGlyHisGlu 180
Db 4988 TACTGGTGACCATGGGACCAACAGTGCACTCAAGAGTGTGACCACACAGGACAGAG 5047
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5048 TTGGCCATGATGACATACCATGGAATTCGGGCTTCTGGCAACCAAGCAATGAAGAA 5107
```

| | | | |
|---|-----------|---|------|
| QY | 201 | GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro | 220 |
| Db | 5108 | GGATGGACACATTTTATGAGTACGACAGCTTTGGCGGCTGACAAATGTGACCTTCCT | 5167 |
| QY | 221 | ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu | 240 |
| Db | 5168 | ACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCAGTGTCCAGGTAGAG | 5227 |
| QY | 241 | ThrSerSerLysAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr | 260 |
| Db | 5228 | ACCTCAGCAAGATGATGTCACCATACCAACCACTGTCGCTCAGCGGCTTCTTAC | 5287 |
| QY | 261 | Thr 261 | |
| Db | 5288 | ACA 5290 | |
| RESULT 4 | | | |
| US-10-383-201-55 | | | |
| ; Sequence 55, Application US/10383201 | | | |
| ; Publication No. US20040029226A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Alsbrook JJ, John et al. | | | |
| ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD | | | |
| ; FILE REFERENCE: 21402-568A | | | |
| ; CURRENT APPLICATION NUMBER: US/10/383,201 | | | |
| ; CURRENT FILING DATE: 2003-03-06 | | | |
| ; PRIOR APPLICATION NUMBER: 10/029020 | | | |
| ; PRIOR FILING DATE: 2001-12-19 | | | |
| ; PRIOR APPLICATION NUMBER: 60/365,984 | | | |
| ; PRIOR FILING DATE: 2002-03-20 | | | |
| ; PRIOR APPLICATION NUMBER: 60/372,022 | | | |
| ; PRIOR FILING DATE: 2002-04-12 | | | |
| ; PRIOR APPLICATION NUMBER: 60/389,143 | | | |
| ; PRIOR FILING DATE: 2002-06-14 | | | |
| ; PRIOR APPLICATION NUMBER: 60/391,779 | | | |
| ; PRIOR FILING DATE: 2002-06-26 | | | |
| ; PRIOR APPLICATION NUMBER: 60/410,755 | | | |
| ; PRIOR FILING DATE: 2002-09-13 | | | |
| ; PRIOR APPLICATION NUMBER: 60/412,957 | | | |
| ; PRIOR FILING DATE: 2002-09-23 | | | |
| ; PRIOR APPLICATION NUMBER: 10/051,874 | | | |
| ; PRIOR FILING DATE: 2002-01-16 | | | |
| ; PRIOR APPLICATION NUMBER: 60/366,928 | | | |
| ; PRIOR FILING DATE: 2002-03-22 | | | |
| ; PRIOR APPLICATION NUMBER: 10/055,877 | | | |
| ; PRIOR FILING DATE: 2002-01-22 | | | |
| ; NUMBER OF SEQ ID NOS: 155 | | | |
| ; SOFTWARE: CuraseqList version 0.1 | | | |
| ; SEQ ID NO 55 | | | |
| ; LENGTH: 8355 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Homo sapiens | | | |
| ; FEATURE: | | | |
| ; NAME/KEY: CDS | | | |
| ; LOCATION: (1)..(8325) | | | |
| US-10-383-201-55 | | | |
| Alignment Scores: | | | |
| Pred. No.: | 9.08e-170 | Length: | 8355 |
| Score: | 1379.00 | Matches: | 260 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 99.64% | Indels: | 0 |
| DB: | 13 | Gaps: | 0 |
| US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-383-201-55 (1-8355) | | | |
| QY | 1 | LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyVala | 20 |
| Db | 4498 | AAGATCAACCGCATCAGGAGGTCCACCTAGTGGAGATCTCATCTGTTGCTGGGGCC | 4557 |
| QY | 21 | ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly | 40 |

APPLICANT: Edinger, Shlomit R
APPLICANT: Ellerman, Karen
APPLICANT: Gunther, Erik
APPLICANT: Smithson, Glenda
APPLICANT: Millet, Isabelle
APPLICANT: MacDougall, John R
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-230
CURRENT APPLICATION NUMBER: US/10/038,854
PRIOR FILING DATE: 2003-01-22
PRIOR APPLICATION NUMBER: 60/258,928
PRIOR FILING DATE: 2000-12-29
PRIOR APPLICATION NUMBER: 60/259,415
PRIOR FILING DATE: 2001-01-02
PRIOR APPLICATION NUMBER: 60/259,785
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: 60/269,814
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/279,832
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,833
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,863
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/283,889
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,447
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/286,683
PRIOR FILING DATE: 2001-04-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 411
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 39
LENGTH: 8473
TYPE: DNA
ORGANISM: Homo sapiens
US-10-038-854-39

Alignment Scores:
Pred. No.: 2,49e-111 Length: 8473
Score: 935.50 Matches: 172
Percent Similarity: 79.01% Conservative: 35
Best Local Similarity: 65.65% Mismatches: 54
Query Match: 67.59% Indels: 1
DB: 17 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-038-854-39 (1-8473)

QY 1 LysileAsnArgileArgInValThrThrSerGlyGluileSerLeuValAlaGlyAla 20
Db 4317 AAAATTAACCGGATAAGCGAGGTACACAGATGCGAAGATCTCTTAGTGGCGGAATA 4376

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4377 CCTTCAGGTGTGACTGCAAAATATGATGCCAATGTGCTGCTTACCAAGATGGAGTGGC 4436

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4437 TAGCCCAAGATGCCAAACTCAGTGGCCCCATCTCCCTGGCTGCTTCCAGATGGTACA 4496

QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgileArgPheIleArgLysAsnLysProphe 80
Db 4497 CTGTATATTCAGATCTAGGGAATATCCGATCGGGCTGTGTCAAGAATAAGCCTTTA 4556

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4557 CTTAACCTATGATCACTTATGAAGTTGGCTGCCCACTCAAGTCAAGACTCTACATCTTT 4616

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4617 GACATCAATGGTACTCACCATAATATCTGTAAGTTTATGTCATCTGGTATTACCTTTACAAT 4676

QY 121 PheThrThrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4677 TTTAGCTACAGCAATGACAATGATATTAATCTGTGTGACAGACAGCAATGGACACCTTT 4736

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 4737 AGAATTTAGACGGGAGCCCAATCGCATGCCAGTTCCAGTGGTGTCTCTCTGATACCAAGTG 4796

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4797 ATATGTTGACAAATAGGAACAATGATGTTTGAAGGCATGCTCTCAAGACTGGAA 4856

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 4857 TTAGTGTGTTTACTTACCATGGAATAGTGGCTTTTAGCCACTAAAAGTGAATGAACT 4916

QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 4917 GGATGGACAACGTTTTTTTGGACTATGACAGTGAAGTCTGTCTGACAAATGTTACGTTTCCA 4976

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 4977 ACTGGAGTGGTCACAAACCTGCATGGGACATGGACAAGGCTATCACAGTGACATTGAG 5036

QY 241 ThrSerSerLys---AspAspValThrThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5037 TCATCTAGCCGAGAAGAAGATGTCAGCATCATCTCAAATCTGTCTCTCGATGATTCCTTC 5096

QY 260 TyrThr 261
Db 5097 TACACC 5102

RESULT 6
US-10-038-854-41
; Sequence 41, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785

201 GlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
4913 GGATGGACACAGTCTTTTGGCTATGACAGTGAAGGTCGTCTGACAAATGTTACGTTCCA 4972
221 ThrGlyGlnValSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
4973 ACTGGAGTGGTCACAAACCTGCATGGGACATGGACAAGGCTATCACAGTGGACATTTGAG 5032
241 ThrSerSerLys---AspAspValThrThrAsnLeuSerAlaSerGlyAlaPhe 259
5033 TCATCTAGCCGAGAGAAGATGTCAGCATCACTTCAATCTCTCTCGATCGATTTCTTC 5092
260 TyrThr 261
5093 TACACC 5098
RESULT 7
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytex, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Molenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Vellizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine B
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennada
; APPLICANT: Millet, Isabelle
; APPLICANT: Macdougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25

PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-41
Alignment Scores:
Pred. No.: 2,49e-111 Length: 8487
Score: 935.50 Matches: 172
Percent Similarity: 79.01% Conservative: 35
Best Local Similarity: 65.65% Mismatches: 54
Query Match: 67.59% Indels: 1
Gaps: 1
DB:
US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-038-854-41 (1-8487)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluLeuSerLeuValAlaGlyAla 20
Db 4313 AAAATTAAACCGGATAGGACAGGTCACACAGATGGAGAAATCTCCTTAGTGGCGGAATA 4372
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4373 CCTTCAGAGTGTGACTGCAAAATATGATGCCAATGTGACTGTGTACCCAGAGTGAGATGGC 4432
QY 41 TyrAlaIleAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4433 TAGCCCAAGATGCCAAACTCAGTGGCCCAATCTCTCCCTGCTGCTTCCAGATGGTACA 4492
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80
Db 4493 CTGTATATTGCAGATCTAGGGAATATCCGGATCGGGCTGTGTCAAGAATAAGCCCTTA 4552
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4553 CTTAACTCTATGAACCTCTATGAAGTTGCGGTCTCCAACTGATCAAGAACTCTACATCTTT 4612
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4613 GACATCAATGGTACTCCACATATACCTAAGTTTAGTCACTGGTGATACCTTTACAAAT 4672
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4673 TTTAGCTACAGCAATGACATGATATTTACTCTGCTGACAGACAGCAATGGCAACCCCTT 4732
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
Db 4733 AGAATTAGACGGGACCCAAATCGCATGCCATTCAGTGGTGTCTCTGATTAACCAAGTG 4792
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4793 ATATGTTGACATAGGAACAAATGATGTTTGAAGGCAATGATGCTCAAGGACTGGAA 4852
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 4853 TTAGTTTGTCTTACTTACCATGGCAATAGTGGCCCTTTTAGCCACTAAAAGTATGAAACT 4912

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 8645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-37

Alignment Scores:

Pred. No.: 2,578-111 Length: 8645
Score: 935.50 Matches: 172
Percent Similarity: 79.01% Conservative: 35
Best Local Similarity: 65.65% Mismatches: 54
Query Match: 67.53% Indels: 1
DB: 17 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-038-854-37 (1-8645)

QY 1 LysilleAsnArgileArgGlnValThrThrSerGlyGluileSerleuValalaglyAla 20
Db 4489 AAAATTAACCGATAGCGAGTGCACACAGATGGAGAAATCTCTTAGTGCCCGGATA 4548
QY 21 ProSerGlyCysAspCysGlyAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4549 CCTTCAGAGTGTGACTGCAAAAATGATGCCAACTGTGACTGTATACCAGAGTCGAGATGC 4608
QY 41 TyzAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4609 TAGCCCAAGAGTCCCAAACTCAGTGCCCACTCTCCCTCCCTGGCTCTCCAGATGTTACA 4668
QY 61 LeuTyrrValAlaAspLeuGlyAsnIleArgPheIleArgLysAsnLysProPhe 80
Db 4669 CTGTATATTGCAGATCTAGGGAATATCCGATCCGGCTGTGTCAAGAAATAGCCCTTTA 4728
QY 81 LeuAsnThrGlnAsnMetTyrrGluLeuSerSerProIleAspGlnGluLeuTyrrLeuPhe 100
Db 4729 CTTAACTCTATGAATCTTCTATGAAGTTCGCTCCCACTGATCAAGAACTCTACATCTTT 4788
QY 101 AspThrThrGlyLysHisLeuTyrrThrClnSerLeuProThrGlyAspTyrrLeuTyrrAsn 120
Db 4789 GACATCAATGGTACTCCCAATATACGTAAATTTAGTCACTGGTGATTTACCTTTTACAAT 4848
QY 121 PheThrTyrrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4849 TTTAGCTACAGCAATGACAAATGATATTACTCTGTGACAGACCAATGGCAACACCTTT 4908
QY 141 AsnValArgArgAspSerThrGlyMetProLeuValValProAspGlyGlnVal 160
Db 4909 AGAATTAGACGGGACCCAAATCGCATGCCAGTTCGAGTGTGTCTCTGTATAACCAAGTG 4968
QY 161 TyrrTrrValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4969 ATATGTTGACATAGGAACAATGGATGTTTGAAGGCGATGCTGCTCAAGGACTGGAA 5028
QY 181 LeuAlaMetMetThrTyrrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5029 TTAGTTTGTGTTTACTTACCATGGCAATAGTGGCTTTTAGCCACTAAAGTGAAGAACT 5088
QY 201 GlyTrrThrThrPheTyrrGluTyrrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5089 GGATGGACAAACGTTTTTTTGACTATGACAGTGAAGTTCGTCTGACAAATGTTACGTTTCCA 5148
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 5149 ACTGGAGTGTGCACAACTCGATGGGACATGGCAAGGCTATCAAGTCGACATTGAG 5208
QY 241 ThrSerSerLys---AspAspValThrIleThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5209 TCATCTAGCCGAGAAGAAGATGTCAGATCACTTCAATCTGCTCCTCGATCGATTTCTTTC 5268
QY 260 TyrrThr 261
|||||

Db 5269 TACACC 5274
RESULT 8
US-10-038-854-35
; Sequence 35, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eissen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35

Alignment Scores:

Pred. No.: 2,588-111 Length: 8675
Score: 935.50 Matches: 172
Percent Similarity: 79.01% Conservative: 35
Best Local Similarity: 65.65% Mismatches: 54
Query Match: 67.53% Indels: 1

```

DB:
US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-038-854-35 (1-8675)
QY 1 LysileAsnArgileArgGlnValThrThrSerGlyGluileSerLeuValalaGlyAla 20
DB 4501 AAATAACCGGATAGGAGGTACACAGATGGAGAAATCTCTTAGTGGCCGGAATA 4560
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspPheSerGlyAspGly 40
DB 4561 CCTTCAGAGTGTGACTGCAAAATGATGCAAACTGTGACTGTACACAGAGTGGAGATGC 4620
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
DB 4621 TAGCGCAAGGATGCCAAATCTAGTGGCCCATCTCTCGCTGGCTCTTCCAGATGGTACA 4680
QY 61 LeuTyrValAlaAspLeuGlyAsnLeuArgileArgPheileArgLysAsnLysProPhe 80
DB 4681 CTTGATATTCAGATCTAGGAAATATCCGATCCGGCTGTGTCAAGAAATAAGCCTTTA 4740
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProileAspGlnGluLeuTyrLeuPhe 100
DB 4741 CTTAACTCTATGAACTCTATGAAGTTGGCTCTCCAACTGATCAAGAACTCTACATCTT 4800
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
DB 4801 GACATCAATGGTACTCACCAATATCTGTAAGTTTACTCTGCTGACAGACAGCAATGCCACCCCTT 4860
QY 121 PheThrTyrThrGlyAspGlyAspPheileThrLeuileThrAspAsnGlyAsnMetVal 140
DB 4861 TTTAGTACAGCAATGACAAATGATTTACTGCTGTGACAGACAGCAATGCCACCCCTT 4920
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
DB 4921 AGAATTAGACGGACCAATCGCATGCCAGTTTCGAGTGGTGTCTCTCGTATAACCAAGTG 4980
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
DB 4981 ATATGGTTGCAATAGGAAACAAATGGATGTTTGAAGGCATGACTGCTCAAGGACTGGAA 5040
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
DB 5041 TTAGTTTGTGTTTACTTACCAATGCAATAGTGGCTTTTAGCCCACTAAAGATGATAACT 5100
QY 201 GlyTrpThrThrPheTyrGlyThrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
DB 5101 GGATGGCAACAGTTTGTGACTATGACAGTCAAGGCTGCTCTGCAAAATGTTACGTTTCCA 5160
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
DB 5161 ACTGGAGTGGTCACAAACCTGCATGGGACATGGACAAAGGCTATCAGATGGACATTTAG 5220
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
DB 5221 TCATCTAGCGGAGAGAGAGTGTGACGATCATCTTCAAATCTGCTCGATCGATCTTTC 5280
QY 260 TyrThr 261
DB 5281 TACACC 5286

RESULT 9
US-10-295-027-927
; Sequence 927, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glyme, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard

```

```

; APPLICANT: Watson, Susan R.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 927
; LENGTH: 12880
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-927

```

```

Alignment Scores:
Pred. No.: 4,16e-106 Length: 12880
Score: 898.00 Matches: 162
Percent Similarity: 78.46% Conservative: 42
Best Local Similarity: 62.31% Mismatches: 56
Query Match: 64.88% Indels: 0
DB: 16 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-295-027-927 (1-12880)
QY 1 LysileAsnArgileArgGlnValThrThrSerGlyGluileSerLeuValalaGlyAla 20
DB 4406 AAAGTAACCGCATTCAGCAAGTAACCAACCAATGGGAGATCTACATCATCGTGGTGCC 4465
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspPheSerGlyAspGly 40
DB 4466 CCCACTGACTGTGACTGCAAAATGATCCAAACTGTGACTGTCTTTTCAGGTGATGGTGGC 4525
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
DB 4526 TATGCCAAAGATGCAAAAGATGAAAGCCCTTCCTCTTAGCAGTGTCCCTGATGGAAACC 4585
QY 61 LeuTyrValAlaAspLeuGlyAsnLeuArgileArgPheileArgLysAsnLysProPhe 80
DB 4586 CTCATGTGGCAGACCTTCGGAAATGTTGGAATTCACCATCAGCAGGAACCAAGCCAC 4645
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProileAspGlnGluLeuTyrLeuPhe 100
DB 4646 CTGAATGACATGAACATTTATGAGATGCTTTCACCGCTGATCAGGAACCTGACAGTTC 4705
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
DB 4706 ACTGTAATGGAACCCACCTACACACCTGAACTTGAATTAACAGGGAATGTTTATTAAC 4765
QY 121 PheThrTyrThrGlyAspGlyAspPheileThrLeuileThrAspAsnGlyAsnMetVal 140
DB 4766 TTCACCTACCAATTCGAAGGTGACTTGGCGCGGATTAACAGCAGCAATGGAATTCAGTG 4825

```


Alignment Scores:

Pred. No.: 4,31e-102 Length: 8689
Score: 865.50 Matches: 163
Percent Similarity: 77.10% Conservative: 39
Best Local Similarity: 62.21% Mismatches: 59
Query Match: 62.54% Indels: 1
DB: 10 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-800-198-66 (1-8689)

QY 1 LysLeuAsnArgIleArgGlnValThrThrSerGlyGluLeuValAlaGlyAla 20
Db 4594 AAGATCAACCGCTACCGAGGTACACCAACGAGAGATCTGCTCTTAGCGGGGCA 4653
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4654 GCCTCAGACTGTGACTGCAAAATGACGTCACTGCATCTGCTATTCGGGAGATGACGCA 4713
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4714 TAGCCACGAGTATCCATCTTGAATCCCTCCCTCTTAGCTGTGGCTCCGGATGGCAAC 4773
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4774 ATCTACATCCAGACCTCGGAGATATCCGATCAGGCGGTACGCAAAACAACTGTT 4833
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4834 CTTAAACGCTTCAACAGTATGAGGTGCTGCTCCGGGAGAACAGGAACCTGTAGCTGTC 4893
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4894 AACCGCATGGTATCCATCATGATGATGACCGCTGTGTCGCGGGAGTACTTATACAA 4953
QY 121 PheThrThrGlyAspGlyAspIleThrLeuLeuThrAspAsnAnglyAsnMetVal 140
Db 4954 TTCACCTACAGCGCTGACATGATGTCACCGAGTGTATTGACAAACACGGGAATTCCTTA 5013
QY 141 AsnValArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 5014 AAGATCCGCGGACAGACAGTGGCATGCCCGACACCTGCTCATGCTGATAATACATG 5073
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5074 ATCACCTTAGTGGGGACCAACGAGGCGCTCAAGACCGGTCAACGACAGCAACTGGAG 5133
QY 181 LeuAlaMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5134 CTGGGCTCATGACTTATGATGGGACACACTGACTCTCTAGCCACCAAGAGCGATGAAAC 5193
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5194 GGATGGACAACTTTTATGACTATGACACAGGCGGCTCTGACCAATGTGACTGCGCC 5253
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 5254 ACGGGGTGGGACCGCTCCACCGGAATGAGAAATCCATCACCCTGTCACATTTAG 5313
QY 241 ThrSerLysAsp---AspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5314 AACTCCACCTGATACGATGTCATCTGATTACCAACCTCTCTTTCAGTGGAGCCCTCC 5373
QY 260 TyrThr 261
Db 5374 TACACC 5379

RESULT 12

US-09-808-602-74
; Sequence 74, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A

; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Heriman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 74
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-808-602-74

Alignment Scores:

Pred. No.: 5,95e-102 Length: 8797
Score: 864.50 Matches: 162
Percent Similarity: 77.10% Conservative: 40
Best Local Similarity: 61.83% Mismatches: 59
Query Match: 62.46% Indels: 1
DB: 9 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-808-602-74 (1-8797)

QY 1 LysLeuAsnArgIleArgGlnValThrThrSerGlyGluLeuValAlaGlyAla 20
Db 4571 AAGATCAACCGCTACCGAGGTACACCAACGAGAGATCTGCTCTTAGCGGGGCG 4630
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4631 GCCTCAGACTGTGACTGCAAAACGATGTCACTGCTACTCGGAGATGAGCT 4690
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4691 TAGCCACGAGCGCATCTTGAATCCGCTGCTCTTAGCGGGCTCCGGATGGCAAC 4750
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4751 ATCTACATTCGAGACCTTGGAGATATCCGATCAGGCGGTACGCAAAATATAACCGCT 4810
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4811 CTTAAACGATTCACACAGTATGAGGTGCTCATCTCCGGGAGAACAGGAATTTGACGTGTC 4870
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4871 AACCTCATGGTATCCATCATGATGATGAGTCTGCTGAGTCTGGGAGTACTTGTACAA 4930
QY 121 PheThrThrGlyAspGlyAspIleThrLeuLeuThrAspAsnAnglyAsnMetVal 140
Db 4931 TTCATATACGCGCTGACATGACGTACCGAGTGTATTGACAAACACGGAATTCCTTA 4990
QY 141 AsnValArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 4991 AAGATCCGCGGAGCAGCAGTGGCATGCGCCGCACTGCTCATGCGCGGAATCAGATT 5050
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5051 ATCACCTTACTGTGGGACCAATGAGGCTCAAGCGGTGCTCACTCAGAACCTGGAG 5110
QY 181 LeuAlaMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5111 CTGGGCTCATGACTTATGATGGGAAACACTGACTCTCTAGCCACCAAGAGTATGAAACC 5170

QY 201 GlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5171 GGATGGACACATTTTATGACTATGACACAGGCGGCTGACCAATGTGACCGCC 5230
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 5231 ACGGGCGTGTGACCACTGACACCGGAAATGGAGAAATCTATCACCATTGACATTGAG 5290
QY 241 ThrSerSerLys---AspAspValThrThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5291 AACTCCACGGGATGATGACGTCATGATGATCACCACCACTCTCTCCGTTGGAGGCTCC 5350
QY 260 TyrThr 261
Db 5351 TATACA 5356
RESULT 13
US-09-808-602-77
; Sequence 77, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-808-602-77
Alignment Scores:
Pred. No.: 5,95e-102 Length: 8797
Score: 864.50 Matches: 162
Percent Similarity: 77.10% Conservative: 40
Best Local Similarity: 61.83% Mismatches: 59
Query Match: 62.46% Indels: 1
DB: 9 Gaps: 1
US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-808-602-77 (1-8797)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4571 AAGATCAACCGCTAGCCCACTGACCACTGAGAGATCTGCTCTTAGCCGGGCG 4630
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4631 GCCTCAGACTGTGACTGCMAAAGCATGTCAACTGCATCTCTACTCGGAGATGACGCT 4690
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerIleAlaValCysAlaAspGlyGlu 60
Db 4691 TACGCCACGGACGCATCTCAACTCGCGCTCTCTTAGCGGTGGCTCCGCGATGGCACC 4750
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4751 ATCTACATTGACACCTTGGGAATATCCGATCAGGCGGTGACGAAAAATAAACCGGTT 4810
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100

Db 4811 CTTAACGCATTTCAACCGATATGAGGCTGCATCTCCGGGAGAAACAGGAATTTGTCGTTC 4870
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4871 AACGCTGATGGTATCCATCAGTACATCTGTGAGTCTGTGACTGGGAGTACTTGTACAAT 4930
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4931 TTCACATACAGGCTGACCAATGACGTCACCGAGTTGATTGACAAACACGCGAATTCCTTA 4990
QY 141 AsnValArgAspAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 4991 AAGATCCGCGGACAGCAGTGGCATGCCCGCCCTCATCTCATCGCGATATATCAGATT 5050
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5051 ATCACCTTTACTGTGGGCACCAATGAGGCGCTCAAGCCGTGTCCACTCAGAACCTGGAG 5110
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5111 CTGGGCTCTATGACTTATGATGGAACTGAGTCTAGCCACCAAGAGTGTGAAACC 5170
QY 201 GlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5171 GGATGGACACATTTTATGACTATGATGACGACGAGGCGGCTGACCAATGTGACCGCC 5230
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 5231 ACGGGCGTGTGACCACTGACCGGGAATGGAGAAATCTATCACCATTGACATTGAG 5290
QY 241 ThrSerSerLys---AspAspValThrThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5291 AACTCCACGGGATGATGACGTCATGATGATCACCACCACTCTCTCCGTTGGAGGCTCC 5350
QY 260 TyrThr 261
Db 5351 TATACA 5356
RESULT 14
US-09-800-198-62
; Sequence 62, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-62
Alignment Scores:
Pred. No.: 5,95e-102 Length: 8797
Score: 864.50 Matches: 162
Percent Similarity: 77.10% Conservative: 40
Best Local Similarity: 61.83% Mismatches: 59
Query Match: 62.46% Indels: 1
DB: 10 Gaps: 1

QY 260 TyThr 261
| | | | |
Db 5351 TATACA 5356

Search completed: August 14, 2004, 19:19:36
Job time : 604.097 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 77.2165 Seconds

(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_1490_1750

Perfect score: 1384

Sequence: 1 KINRIRQVTSIGISLVAG.....SSKDDVTITNLSASGAFYT 261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=Issued Patents NA -QEMT=fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn 1.1.258 @runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 124.5 | 9.0 | 4881 | 4 | US-09-252-991A-7234 |
| 2 | 115 | 8.3 | 31960 | 4 | US-09-453-702B-11 |
| 3 | 112.5 | 8.1 | 2415 | 4 | US-09-328-352-1419 |
| 4 | 105.5 | 7.6 | 4761 | 4 | US-09-543-681A-1157 |
| 5 | 105 | 7.6 | 640681 | 4 | US-09-790-998-1 |
| 6 | 101.5 | 7.3 | 885 | 4 | US-09-446-301A-3 |
| 7 | 101.5 | 7.3 | 885 | 4 | US-09-099-932-3 |
| 8 | 101.5 | 7.3 | 1607 | 4 | US-09-446-301A-16 |
| 9 | 101.5 | 7.3 | 1607 | 4 | US-09-099-932-50 |
| 10 | 98.5 | 7.1 | 1515 | 4 | US-09-252-991A-11708 |
| 11 | 98 | 7.1 | 4860 | 4 | US-09-328-352-3221 |
| 12 | 97.5 | 7.0 | 1260 | 4 | US-09-252-991A-15163 |

| | | | | | | |
|------|------|-----|-------|---|----------------------|-------------------|
| C 13 | 97.5 | 7.0 | 2181 | 4 | US-09-252-991A-14909 | Sequence 14909, A |
| C 14 | 97.5 | 7.0 | 3387 | 4 | US-09-252-991A-14461 | Sequence 14461, A |
| C 15 | 96.5 | 7.0 | 2236 | 4 | US-09-221-017B-482 | Sequence 482, Ap |
| C 16 | 96 | 6.9 | 3015 | 4 | US-09-206-942-56 | Sequence 56, Appl |
| C 17 | 96 | 6.9 | 3033 | 4 | US-09-206-942-54 | Sequence 54, Appl |
| C 18 | 94.5 | 6.8 | 11613 | 4 | US-09-453-702B-42 | Sequence 42, Appl |
| C 19 | 94 | 6.8 | 642 | 4 | US-09-252-991A-7031 | Sequence 7031, Ap |
| C 20 | 94 | 6.8 | 684 | 4 | US-09-252-991A-11528 | Sequence 11528, A |
| C 21 | 94 | 6.8 | 1593 | 4 | US-09-252-991A-7095 | Sequence 7095, Ap |
| C 22 | 94 | 6.8 | 2715 | 4 | US-09-543-681A-2771 | Sequence 2771, Ap |
| C 23 | 94 | 6.8 | 4453 | 4 | US-09-453-702B-90 | Sequence 90, Appl |
| C 24 | 93.5 | 6.8 | 795 | 4 | US-09-134-001C-2198 | Sequence 2198, Ap |
| C 25 | 93.5 | 6.8 | 804 | 4 | US-09-134-001C-1702 | Sequence 1702, Ap |
| C 26 | 93 | 6.7 | 3900 | 4 | US-09-023-655-1420 | Sequence 1420, Ap |
| C 27 | 93 | 6.7 | 11935 | 4 | US-09-634-238-401 | Sequence 401, Ap |
| C 28 | 92.5 | 6.7 | 1785 | 4 | US-09-489-039A-6762 | Sequence 6762, Ap |
| C 29 | 92 | 6.6 | 5319 | 1 | US-08-169-927-1 | Sequence 1, Appl |
| C 30 | 92 | 6.6 | 25165 | 4 | US-09-453-702B-39 | Sequence 39, Appl |
| C 31 | 91.5 | 6.6 | 2781 | 4 | US-09-313-677-1 | Sequence 1, Appl |
| C 32 | 91.5 | 6.6 | 2799 | 4 | US-09-313-677-18 | Sequence 18, Appl |
| C 33 | 91.5 | 6.6 | 7026 | 4 | US-09-313-677-20 | Sequence 20, Appl |
| C 34 | 91.5 | 6.6 | 7344 | 4 | US-09-313-677-16 | Sequence 16, Appl |
| C 35 | 91 | 6.6 | 49617 | 4 | US-09-596-002-28 | Sequence 28, Appl |
| C 36 | 90.5 | 6.5 | 5738 | 1 | US-08-409-995-3 | Sequence 3, Appl |
| C 37 | 90.5 | 6.5 | 5738 | 3 | US-08-685-467-3 | Sequence 3, Appl |
| C 38 | 90.5 | 6.5 | 7291 | 3 | US-08-913-942-3 | Sequence 3, Appl |
| C 39 | 89.5 | 6.5 | 11679 | 4 | US-09-328-352-1377 | Sequence 1377, Ap |
| C 40 | 89 | 6.4 | 1443 | 4 | US-09-724-623-15 | Sequence 15, Appl |
| C 41 | 89 | 6.4 | 1683 | 4 | US-09-252-991A-6977 | Sequence 6977, Ap |
| C 42 | 89 | 6.4 | 1683 | 4 | US-09-252-991A-11588 | Sequence 11588, A |
| C 43 | 89 | 6.4 | 2925 | 4 | US-09-252-991A-7069 | Sequence 7069, Ap |
| C 44 | 89 | 6.4 | 2997 | 4 | US-09-252-991A-11853 | Sequence 11853, A |
| C 45 | 88.5 | 6.4 | 2745 | 4 | US-09-817-514A-5 | Sequence 5, Appl |

ALIGNMENTS

RESULT 1

US-09-252-991A-7234

; Sequence 7234, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 7234

; LENGTH: 4881

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-7234

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred No.: | 6.01e-05 | Length: | 4881 |
| Score: | 124.50 | Matches: | 53 |
| Percent Similarity: | 41.33% | Conservative: | 40 |
| Best Local Similarity: | 23.56% | Mismatches: | 98 |
| Query Match: | 9.00% | Indels: | 35 |
| DB: | 4 | Gaps: | 7 |

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-252-991A-7234 (1-4881)

QY 51 SerSertLeuAlaValCysAlaAspGly-----GlutLeuTyValAlaAspLeuGlyAsn 68

Db 2143 AGCGTTACCGCATCAACGCCGACGCGAGGAGTCTATGTCCAGAC----- 2193

QY 69 lleArglleAcpPheilleArglyAsnLysPro---PheLeuAsnThrGlnAsnMetTyr 87
Db 2194 GACAAACGGCGCTGTCAGCGAGTGCATCCGAGCGAGGGAACCCCTCAGGCATAC 2253
QY 88 GlnLeuSerSerProilleAspGlnGluLeu-----TyrLeu 99
Db 2254 GACGAGAAGGGCACTGTCGCCGAAACCGATCCGTTGGGAACGATCACCGAGTACCGC 2313
QY 100 PheAspThrThrGlyLys----- 105
Db 2314 TACGACCGCGCGAGCTTGGAGGCACTGCTCCCGCGGAGCGAGCAACCTGCTAT 2373
QY 106 -----HisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsnPheThr 122
Db 2374 AGCTACTTCGACGGCTTGTGCGCAGCGTGGCGCGCGAGCGCGAGTGGAGTACGAG 2433
QY 123 TyrThrGlyAspGlyAspilleThrLeuileThrAspAsnAsnGlyAsnMetValAsnVal 142
Db 2434 CGCAACGCTCAGGGCGACATCACCCGCCAGACCGATCCCGAAGGCAACGTCACCCACTAC 2493
QY 143 ArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnValTyrTrp 162
Db 2494 GCCTACGACACCGCGCTGCTGCTGGTGAATCACCCATGCCGACGCGAGCCCTGACCCAG 2553
QY 163 valThrMetGlyThrAsnSerAlaLeuLysSerValThrGlnGlyHisGluLeuAla 182
Db 2554 TTGACCTGG-----AATCGCTCGCGCATTTGATAGAACACAGCTTCGCGAGGTAGC 2607
QY 183 MetMetThrTyr---HisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsnGly 201
Db 2608 GTGCGCGCTATCGCTACGACACACTGGGTAGACAGATCACCCGCCAGGAGCAAGCGCG 2667
QY 202 TrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPheProThr 221
Db 2668 GCCATCACCGGTTCCAAATGGATGTCAGACGCGCGCTGAGCCAGATCACCCCTGCCCGGT 2727
QY 222 GlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGluThr 241
Db 2728 GCGCGCAGCGCGATGATCCCTACACGCTTACGGCAAGGT-CACCGAGCGAGTGCAGCA 2786
QY 242 SerSerLysAspAsp 246
Db 2787 ACAGGCGCGGATGAC 2801

RESULT 2
US-09-453-702B-11
; Sequence 11, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 31960
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-453-702B-11
Alignment Scores:
Pred. No.: 0.0169 Length: 31960
Score: 115.00 Matches: 46
Percent Similarity: 31.73% Conservative: 20
Best Local Similarity: 22.12% Mismatches: 72
Query Match: 8.31% Indels: 70
Gaps: 5
DB:

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-453-702B-11 (1-31960)
QY 108 TyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsnPheThrTyrThrGlyAspGly 127
Db 26461 TACCCGGATGAATGCGCGCGCGCGCG---CTGCGCGCTACACGTACACGCGCGCGG 26517
QY 128 AspilleThrLeuileThrAspAsnAsnGlyAsnMetVal----- 140
Db 26518 GAATCGGCGCGGTGATGACCGCGAGCGGACCGAGGTGCGCGGTTGCTTATGATGCG 26577
QY 141 -----AsnValArg 143
Db 26578 GAGCAGCGCGCGGATGTTGGCGCACATTATCGCGGTAGCGCGGAGAGCGCTACCGG 26637
QY 144 ArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnValTyrTrpVal 163
Db 26638 TATGATGATACCGCGCGCGGTGACGAGCTGTCUACCCGAGGCGGTGACTACCGCTTT 26697
QY 164 ThrMetGlyThrAsn----- 168
Db 26698 GAGTACGGCGAGGACCGCTGTGACCATCACGAGACCGCTGAACCGCGGAGGTCTGTAC 26757
QY 169 -----SerAlaLeuLysSerValThrGlnGlyHisGluLeuAlaMetMet 184
Db 26758 ACGGAAGCGAGGTTGGCTGAACGCTGTTGTGAAGAAGAACATCGGAGCGGAGCATC 26817
QY 185 Thr----- 185
Db 26818 ACCCGAGCGAGTATGATGAGCGGAGGCTGAGGACACAGACGATCGCGCGGAGCGG 26877
QY 186 -----TyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 26878 CGGACGGAGTACACCTGCATATCGCGTGGGTGCGGTGACAGCGGTGACGCGGCGCGG 26937
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 26938 GGCAGGACGTTGCGGTATGCTATAACAGCGCGGCGGAGGTGAGCTAGTACGAGTACCG 26997
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 26998 GACGCGTGCAGCAGC-----CGGAGTATGATGAGAAGGAGGAGCTGCGCGCGGAG 27051
QY 241 ThrSerSerLysAspValThr 248
Db 27052 ACCTCGCGCAGCGGAGAGACGACG 27075

```
RESULT 3
US-09-328-352-1419
; Sequence 1419, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1419
; LENGTH: 2415
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1419

Alignment Scores:
Pred. No.: 0.000664 Length: 2415
Score: 112.50 Matches: 47
Percent Similarity: 36.11% Conservative: 16
Best Local Similarity: 26.11% Mismatches: 81
Query Match: 8.13% Indels: 35
DB: 4 Gaps: 5

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-328-352-1419 (1-2415)
QY 91 SerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLeuTyrThrGln 110
Db 767 AGTCCTGTGGAGAAAACCTGGTG-----CAATTGTACAAACAA 805
QY 111 SerLeuProThrGlyAspTyrLeuTyrAsnPheThrThrGlyAspGlyAspIleThr 130
Db 806 CAGATAGCTTT-GGTACACAGATCTAATTATTAGCAGCAGGAGATTTTGTGCTGAT 864
QY 131 LeuIleThrAspAsnAsnGlyAsnMetValAsnValArgAspSerThrGlyMetPro 150
Db 865 TTAAAAACA-----CTGTTGCACGTAAGGTACA-----894
QY 151 LeuTrpLeuValProAspGlyGlnValTyrTrpValThrMetGlyThrAsnSerAla 170
Db 895 -----GATGGATCAGTCTTTCAGACAACTGAATTTGAAAACGCA 936
QY 171 Leu-----LysSerValThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGly 188
Db 937 CTTTAAAAACCAAAAGTTGCGGTATCGGTAGTCTACACAAAGTTATACATATAACTCT 996
QY 189 AsnSerGlyLeuLeuAlaThrLysSerAsnGluAsnGlyTyrThrThrPheTyrGluTyr 208
Db 997 AATACAACTTTAGCAACAGAAAAAGATAATAGTATTAGTGGCAAAAGACATTTAAATAT 1056
QY 209 AspSerPheGlyArgLeuThrAsnValThrPheProThrGlyGlnValSerSerPheArg 228
Db 1057 GATGATACGCGCGGAATTACTTCAATTACCATCCAGATAGTTTCAGTAGAAACGATCAA 1116
QY 229 -----SerAspThrAspSerSerValHisValGlnVal 239
Db 1117 TATTTCAATAAAGGACCTCATCGCTCCGGACATGGCGAGAGTTTGACAACTTAT 1176
QY 240 GluThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 1177 AGTTATTCTCTGTGGCGGTTTAAACAAACGACTAATGCAATATATTAGTGAAGCAATT 1236

RESULT 4
US-09-543-681A-1157
; Sequence 1157, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABII
; FILE REFERENCE: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
```

```
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1157
; LENGTH: 4761
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (4726)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-543-681A-1157

Alignment Scores:
Pred. No.: 0.0143 Length: 4761
Score: 105.50 Matches: 44
Percent Similarity: 36.00% Conservative: 28
Best Local Similarity: 22.00% Mismatches: 73
Query Match: 7.62% Indels: 55
DB: 4 Gaps: 7

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-543-681A-1157 (1-4761)
QY 85 AsnMetTyrGluLeuSerSerProIleAspGlnGlu-----LeuTyrLeuPheAspThrThr 103
Db 2125 AATTTAACCAATGACGCGATCCGGAACAACAATCGACTTCTATGATTGGCAAAAGAT 2184
QY 104 GlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsnPheThrTyr 123
Db 2185 TTTGCGTTTCCACGGCACAACCTTGCCAAATGGG---GCGGCTTGGCATTTGGGAGTAC 2241
QY 124 ThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetValAsnVal--- 142
Db 2242 AATGAGCAGCGCGATATTCGTCGTGTGATTGACCCGCTTGCCCATATCAGCGCTTGCG 2301
QY 142 -----142
Db 2302 TGGGATGACCAAGGCTGTGTCTGGGCAAGTGGATGCTAAGGGTAATGAACCCACTAT 2361
QY 143 -----ArgArgAspSerThrGlyMetProLeuTrp 152
Db 2362 CGCTATAATGCCCGGTGCTAGTTAATCGAGCAACGGGACTGTTCCGGTTATCCCAACCAC 2421
QY 153 LeuValValProAspGlyGlnValTyrTrpValThrMetGlyThrAsnSerAlaLeuLys 172
Db 2422 TTGACCTACGATGAT-----TGG-----GGCAACTTCGC 2451
QY 173 SerValThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeu 192
Db 2452 TCGTTGACCAATGCACAGAATGAACACGACTTACACCTTT---AGTGAAGCGGGTTG 2508
QY 193 LeuAlaThrLysSerAsnGluAsnGlyTyrThrThrPheTyrGluTyrAspSerPheGly 212
Db 2509 TTGTTAACAGTGCCTTACCGGATGGGACAGAAAACGTTTATGACTAATGATGACCGCGG 2568
QY 213 ArgLeuThrAsnValThrPhePro-----Thr 221
Db 2569 CAATTAGTGGGATAACGGATGCGGAGAGCGCCATATTCGTTCGCCCGTAACCGTCGT 2628
QY 222 GlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGluThr 241
Db 2629 GGGCAAGTGTATAGCCCGACGCGATCCGCGAGGCGCATTTGGTTGCAATTTTCTATGACT 2688

RESULT 5
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEKI
; APPLICANT: HATTORI, MASAHIRA
```


APPLICANT: Allignet, Jeanine
TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
TITLE OF INVENTION: COMPOUNDS
FILE REFERENCE: 03495.0173-00000
CURRENT APPLICATION NUMBER: US/09/099,932
CURRENT FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/050,380
EARLIER FILING DATE: 1997-06-20
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 885
TYPE: DNA
ORGANISM: Staphylococcus
US-09-099-932-3

Alignment Scores:
Pred. No.: 0.00346 Length: 885
Score: 101.50 Matches: 49
Percent Similarity: 41.45% Conservative: 31
Best Local Similarity: 25.39% Mismatches: 83
Query Match: 7.33% Indels: 30
DB: 4 Gaps: 10

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-099-932-3 (1-885)

```
QY 68 AsnIleArgIleAArgPheIleAArgLysAsnLysProPheLeuAsnThr-GlnAsnMetTy 87
DB 48 TCCATACGGTATAACTTCATCAGAAGACGGAAGGTATGGTTCCACAACTAAGGCAAA 107
QY 87 rGluLeuSerSerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLe 107
DB 108 CAAATCAGCAGT-----CTAGATCAGAGTGGTAGG---AT 140
QY 107 uTyrThrGlnSerLeuProThrGlyAsp---TyrLeuTyrAsnPheThrTyrThrGlyAs 126
DB 141 AAAAGAATTCGAAGTCTCTACCCCTGATGCTAAAGTGATGTTTAATGTATCTTCACT 200
QY 126 pGlyAspIleThrLeuLeThrAspAsnAsnGlyAsnMetValAsnValArgA-gaspSe 146
DB 201 TGGAGACATA---TGGTTTACAGAGATGGTGCAATAAATAAATCGGAAGCTCTCAAAAAA 257
QY 146 rThrGlyMetProLeuTrp---LeuValValProAspGlyGlnValTyrTrpValThrMe 165
DB 258 AGTGGCTTTACAGAAATCAATGCCACACCGGATTCGTGCTCTACGGAATAACGGA 317
QY 165 tGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGluLeuAlaMetMetTh 185
DB 318 AGGTCTAAATGGCGATATATGGTTTACCAATTCGAATGGAGATCGTATAGGAAGTTGAC 377
QY 185 rTyrHisGly-----AsnSerGly-----LeuLeuAl 194
DB 378 AGCTGATGGGACTATTTATGAATATGATTTGCCAAATAAGGACTTATCTGCTTTTAT 437
QY 194 aThrLysSerAsnGluAsnGly---TTPThrThrPheTyrGluTyrAspSerPheGlyAr 213
DB 438 TACTTTAGTTCGGATAACGCACTTTGTTCCAGGAGAACCAAAATAATCTATTATGGAAG 497
QY 213 gLeuThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAspThrAspSe 233
DB 498 GATTACAAAT-----ACAGGAAATTAGAGAATATCTCTACCAACAATATGC 545
QY 233 rSerValHisValGlnValGluThrSerSerLysAsp 245
DB 546 AGCGGCTCCAGTGGGTATC---ACTAGTGTAAACGAT 579
```

RESULT 8

US-09-446-301A-16
Sequence 16, Application US/09446301A
Patent No. 6506893
GENERAL INFORMATION:
APPLICANT: EL SOLH, NEVINE

APPLICANT: ALLIGNET, JEANINE
TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
TITLE OF INVENTION: COMPOUNDS
FILE REFERENCE: 03715-0059
CURRENT APPLICATION NUMBER: US/09/446,301A
CURRENT FILING DATE: 1999-12-20
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 1607
TYPE: DNA
ORGANISM: Staphylococcus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (39..923, 947..1582)
US-09-446-301A-16

Alignment Scores:
Pred. No.: 0.00863 Length: 1607
Score: 101.50 Matches: 49
Percent Similarity: 41.45% Conservative: 31
Best Local Similarity: 25.39% Mismatches: 83
Query Match: 7.33% Indels: 30
DB: 4 Gaps: 10

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-446-301A-16 (1-1607)

```
QY 68 AsnIleArgIleAArgPheIleAArgLysAsnLysProPheLeuAsnThr-GlnAsnMetTy 87
DB 86 TCCATACGGTATAACTTCATCAGAAGACGGAAGGTATGGTTCCACAACTAAGGCAAA 145
QY 87 rGluLeuSerSerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLe 107
DB 146 CAAATCAGCAGT-----CTAGATCAGAGTGGTAGG---AT 178
QY 107 uTyrThrGlnSerLeuProThrGlyAsp---TyrLeuTyrAsnPheThrTyrThrGlyAs 126
DB 179 AAAAGAATTCGAAGTCTCTACCCCTGATGCTAAAGTGATGTTTAATGTATCTTCACT 238
QY 126 pGlyAspIleThrLeuLeThrAspAsnAsnGlyAsnMetValAsnValArgA-gaspSe 146
DB 239 TGGAGACATA---TGGTTTACAGAGATGGTGCAATAAATAAATCGGAAGCTCTCAAAAAA 295
QY 146 rThrGlyMetProLeuTrp---LeuValValProAspGlyGlnValTyrTrpValThrMe 165
DB 296 AGTGGCTTTACAGAAATCAATGCCACACCGGATTCGTGCTCTACGGAATAACGGA 355
QY 165 tGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGluLeuAlaMetMetTh 185
DB 356 AGGTCTAAATGGCGATATATGGTTTACCAATTCGAATGGAGATCGTATAGGAAGTTGAC 415
QY 185 rTyrHisGly-----AsnSerGly-----LeuLeuAl 194
DB 416 AGCTGATGGGACTATTTATGAATATGATTTGCCAAATAAGGACTTATCTGCTTTTAT 475
QY 194 aThrLysSerAsnGluAsnGly---TTPThrThrPheTyrGluTyrAspSerPheGlyAr 213
DB 476 TACTTTAGTTCGGATAACGCACTTTGTTCCAGGAGAACCAAAATAATCTATTATGGAAG 535
QY 213 gLeuThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAspThrAspSe 233
DB 536 GATTACAAAT-----ACAGGAAATTAGAGAATATCTCTACCAACAATATGC 583
QY 233 rSerValHisValGlnValGluThrSerSerLysAsp 245
DB 584 AGCGGCTCCAGTGGGTATC---ACTAGTGTAAACGAT 617
```

RESULT 9

US-09-099-932-50
Sequence 50, Application US/09099932
Patent No. 6570001
GENERAL INFORMATION:

```

; APPLICANT: El Solh, Nevine
; APPLICANT: Allignet, Jeanine
; TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
; TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
; TITLE OF INVENTION: COMPOUNDS
; FILE REFERENCE: 03495.0173-00000
; CURRENT APPLICATION NUMBER: US/09/099,932
; CURRENT FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/050,380
; EARLIER FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 1607
; TYPE: DNA
; ORGANISM: Staphylococcus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(923)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (947)..(1582)
; US-09-099-932-50

```

```

Alignment Scores:
Pred. No.: 0.00863      Length: 1607
Score: 101.50
Percent Similarity: 41.45%
Conservative: 49
Best Local Similarity: 25.39%
Mismatch: 83
Query Match: 30
DB: 10
Gaps: 10

```

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-099-932-50 (1-1607)

```

QY 68 AsnileAargPheileAargLysAsnLysProPheLeuAsnThr-GlnAsnMetTy 87
Db 86 TCATACCGTATAAATTCATCAGAACGCGAAAGGTATGTTTCACACATAAAGCAAA 145
QY 87 rgluLeuSerSerProileAspGlnGluLeuTyLeuPheAspThr-ThrGlyLysHisLe 107
Db 146 CAAATCAGCAGT-----CTAGATCAGAGTGGTAGG---AT 178
QY 107 uTyThrGlnSerLeuProThrGlyAsp---TyrLeuTyAsnPheThrTyThrGlyAs 126
Db 179 AAAAGAAATCGAAGTTCCTACCTGATGCTAAGATGATGTTAATTTGATCTTCACT 238
QY 126 pGlyAspilleThrLeulleThrAspAsnAsnGlyAsnMetValAsnValargargaspSe 146
Db 239 TGGAGACATA---TGGTTTACAGAGAATGCTGCAAAATAAAATCGAAAGCTCTCAAAAAA 295
QY 146 rThrGlyMetProLeuTrip---LeuValValProAspGlyGlnValTyTripValThrMe 165
Db 296 AGTGCGTTTACAGAAATATCCATTCACACAGCGGATTCGGTCCCTACGGAATACGGA 355
QY 165 tGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisLeuLeuAlaMetMetTh 185
Db 356 AGTCTAAATGGCGATATATGTTTACCCCAATGAATGGAGATCGTATAGGAAGTTGAC 415
QY 185 rTyHisGly-----AsnSerGly-----LeuLeuAl 194
Db 416 AGCTGATGGGACTATTATGATATATGATTTGCCAAATAAAGGATCTTATCCTGCTTTAT 475
QY 194 aThrLysSerAsnGluAsnGly---TrpThrThrPheTyGluTyAspSerPheGlyAr 213
Db 476 TACTTTAGTTCGATACCACTTTGGTTCACGAGAACCAAAATTAATTTCTATTGGAAG 535
QY 213 gLeuThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAspThrAspSe 233
Db 536 GATTACAAAT-----ACAGGCAAAATTAGAAGAATATCTCTACCAACAAATATGC 583
QY 233 rSerValHisValGlnValGluThrSerSerLysAsp 245
Db 584 AGCGGCTCCAGTGGGTATC---ACTAGTGTAAAGCAT 617

```

RESULT 10

```

US-09-252-991A-11708
; Sequence 11708, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11708
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-11708

```

```

Alignment Scores:
Pred. No.: 0.0188      Length: 1515
Score: 98.50
Percent Similarity: 37.20%
Conservative: 35
Best Local Similarity: 23.20%
Mismatch: 104
Query Match: 7.12%
DB: 4
Gaps: 12

```

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-252-991A-11708 (1-1515)

```

QY 10 ThrSerGlyGluileSerLeuValAlaGlyAlaProSerGlyCysAspCysLysAsnAsp 29
Db 274 ACCTCGGGAGCTACCTTGATCGGCAACAGCTGCCACTAGGGCTGAGCTTATAACGAG 333
QY 30 ---AlaAsnCysAspCysPheSerGly-----AspAspGlyTyAlaLysAspAla 45
Db 334 ACAGGAAGCTGCGATTGTCGCCCGCGTTACGAAATGAATGATCGCTGGCTGAAGCTCCA 393
QY 46 LysLeuAsnThrProSerSerLeuAlaVal-----CysAlaAspGlyGluLeu 61
Db 394 GATAAAATATGGCCCCCAGCAAGGGGGCTCCGCCCGAGGGTTGCCGCGAAACCCCGGTC 453
QY 62 TyValAlaAspLeuGlyValAsnIleAArgPheileAArgLysAsnLysProPheLeu 81
Db 454 AACATCACCAT---GGCAAC-----471
QY 82 AsnThrGlnAsnMetTyThrGluLeuSerSerProileAspGlnGluLeuTyLeu----- 99
Db 472 AAATATCAGTCCGAGCATGACTTAATAACCCCGATTCGGCTTCCCGCACACTACAATGGC 531
QY 100 PheAspThrThrGlyLysHisLeuTyThrGlnSerLeuProThrGlyAspTyLeuTyTr 119
Db 532 CTGCGCGAATATGCGGCTCATTCGTTCTCTGCGCGCATACACGCAAGGATGATAGCTAC 591
QY 120 AsnPheThrTyThrGlyAspGly-----AspIle 129
Db 592 ---CTATTTCGCGAAGACGGAAGGTTTCGGAGTTTCAGTGGGGCGCGGAGAGATCTA 648
QY 130 ThrLeuIleThrAsp-----AsnAsn 136
Db 649 ACTTCCTCAGCGATCTGGGAAAGCTTAGCCGACTAGCAGGAAGATTTTCTATACCTCG 708
QY 137 GlyAsnMetValAsnValArgAspSerThrGlyMetProLeuTyLeuValValPro 156
Db 709 GAGCTCAACGAATACTATAGAGTTTGACCCCTATGGAATACTCGCCAGCTCAAGACCAAG 768
QY 157 AspGlyGlnValTyTripValThrMetGlyThrAsnSerAlaLeuLysSerValThrThr 176
Db 769 GAGGTCGAAGATATCGTGTAGAGCGAGCGCTAATCTTACGATTAGCGAC-----GAA 822

```



```
Db 736 GCGAGCGCCCTGAGCTATGATACAGACGCCCTTGGGGCGCAGACCGGTATC-----GCC 683
QY 175 ThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAla 194
Db 682 GATGCGGAGGGCAGCCACGCTGTTTCTCGGGGCGACGGCGAT-----CTGCTGGCG 629
QY 195 ThrLysSerAsnGluAsnGlyTrpThrPheTyrGluTyrAspSerPheGlyArgLeu 214
Db 628 CGGGTTTCCGAGCGCGCGCGCGGCGGAAATGTCTCTACCTGCACGACGAACCGCGAGGCTG 569
QY 215 ThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAsp 230
Db 568 GTCGCCCTGACCAACAGAGACGGCGTCCAGGCGCAGTTCGGCTACGAC 521

RESULT 13
US-09-252-991A-14909/c
; Sequence 14909, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14909
; LENGTH: 2181
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14909

Alignment Scores:
Pred. No.: 0.044 Length: 2181
Score: 97.50 Matches: 30
Percent Similarity: 42.24% Conservative: 19
Best Local Similarity: 25.86% Mismatches: 60
Query Match: 7.04% Indels: 7
DB: 4 Gaps: 3

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-252-991A-14909 (1-2181)
QY 118 LeuTyrAsnPheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGly 137
Db 678 GTCTGGCGTTACGACTACGACAAACGAGGACGTTTGTGCGCCAGCAGCAGCCCGCTGGGA 619
QY 138 AsnMetValAsnValArgAspSerThrGlyMetProLeuTyrLeuValProAsp 157
Db 618 CAACTGACCGCGCGCGCTACGACCCGCTGGGCACTGATCGCCCTGGAAATGGCCGAC 559
QY 158 GlyGlnValTyr-----TrpValThrMetGlyThrAsnSerAlaLeuLysSerVal 174
Db 558 GGCAGCGCCCTGAGCTATGATACGACGCCCTTGGGCGGCGACACCGCTATC-----GCC 505
QY 175 ThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAla 194
Db 504 GATGCCGAGGGCGACGCCACGCTGTTTCTCGGGGCGACCGCGAT-----CTGCTGGCG 451
QY 195 ThrLysSerAsnGluAsnGlyTrpThrPheTyrGluTyrAspSerPheGlyArgLeu 214
Db 450 CGGGTTTCCGAGCGCGCGCGCGGCGAATGTCCTACCTGCACGACGAACCGCGAGGCTG 391
QY 215 ThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAsp 230
Db 390 GTCGCCCTGACCAACAGAGACGGCGTCCAGGCGCAGTTCGGCTACGAC 343

RESULT 14
US-09-252-991A-14461
; Sequence 14461, Application US/09252991A
```

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14461
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14461

Alignment Scores:
Pred. No.: 0.0864 Length: 3387
Score: 97.50 Matches: 30
Percent Similarity: 42.24% Conservative: 19
Best Local Similarity: 25.86% Mismatches: 60
Query Match: 7.04% Indels: 7
DB: 4 Gaps: 3

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-252-991A-14461 (1-3387)
QY 118 LeuTyrAsnPheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGly 137
Db 1513 GTCTGGCGTTACGACTACGACAAACGAGGACGTTTGTGCGCCAGCAGCAGCCCGCTGGGA 1572
QY 138 AsnMetValAsnValArgAspSerThrGlyMetProLeuTyrLeuValProAsp 157
Db 1573 CAACTGACCGCGCGCGCTACGACCCGCTGGGCACTGATCGCCCTGGAAATGGCCGAC 1632
QY 158 GlyGlnValTyr-----TrpValThrMetGlyThrAsnSerAlaLeuLysSerVal 174
Db 1633 GGCAGCGCCCTGAGCTATGATACGACGCCCTTGGGCGGCGACACCGCTATC-----GCC 1686
QY 175 ThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAla 194
Db 1687 GATGCCGAGGGCGACGCCACGCTGTTTCTCGGGGCGACCGCGAT-----CTGCTGGCG 1740
QY 195 ThrLysSerAsnGluAsnGlyTrpThrPheTyrGluTyrAspSerPheGlyArgLeu 214
Db 1741 CGGGTTTCCGAGCGCGCGCGCGGCGAATGTCCTACCTGCACGACGAACCGCGAGGCTG 1800
QY 215 ThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAsp 230
Db 1801 GTCGCCCTGACCAACAGAGACGGCGTCCAGGCGCAGTTCGGCTACGAC 1848

RESULT 15
US-09-221-017B-482/c
; Sequence 482, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
```



```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/221,017B
/ FILING DATE: 23-DEC-1998
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PP1182
/ FILING DATE: 31-DEC-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PP1546
/ FILING DATE: 30-JAN-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PP2911
/ FILING DATE: 09-APR-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/AU98/01023
/ FILING DATE: 10-DEC-1998
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Monroy, Gladys H
/ REGISTRATION NUMBER: 32,430
/ REFERENCE/DOCKET NUMBER: 27340-20021.00
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-813-5600
/ TELEFAX: 650-494-0792
/ TELEX: 706141
/ INFORMATION FOR SEQ ID NO: 482:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2236 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: UNKNOWN
/ ORIGINAL SOURCE:
/ ORGANISM: PORYPHYROMONAS GINGIVALIS
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 1...2236
US-09-221-017B-482
```

```
Alignment Scores:
Pred. No.: 0.0611 Length: 2236
Score: 96.50 Matches: 65
Percent Similarity: 34.12% Conservative: 22
Best Local Similarity: 25.49% Mismatches: 85
Query Match: 6.97% Indels: 83
DB: 4 Gaps: 14

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-221-017B-482 (1-2236)

Qy 47 LeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGluLeuTyrValAlaAspLeu 66
Db 1907 TTGGAAGCTCCCCACCCTTCGGGTG---GTCGAAGGTATCGTAGAATGGCTGATCTC 1851
Qy 67 GlyAsnIleArgIleArgPhe-----IleArg 75
Db 1850 GGCAGAAATTGCGAAGCGGTCTCTTCGTAATTAGGTGATACGCTCCACAGCATTAGA 1791
Qy 76 LysAsnLysProPheLeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGln 95
Db 1790 GCCGAACAGCCGGAACCTCCACCGCGAATTACTTC-----1755
Qy 96 GluLeuTyrLeuPheAspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGly 115
Db 1754 -----TACAGTTTCGATCATATTGGCAGGATCTG---CTCCAAGCCATAGACACCGGC 1704
Qy 116 AspTyrLeuTyrAsnPheThrTyrThrGlyAspIleThrLeuIleThrAspAsn 135
Db 1703 -----AAGCGAACTGAAGATGG-----ATGGCTGTCGATAAG 1671
Qy 136 Asn-----GlyAsnMetValAsnValArgAspSerThrGlyMetProLeuTyr 152
Db 1670 AATTTCGAATAGGCTCCTTCGAGTCCATTGATACGAACCTGGTTGAAACCGCAGTTCTG 1611
```

Search completed: August 14, 2004, 21:37:30
Job time : 323.216 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 3060.96 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_1490_1750
Perfect score: 1384
Sequence: 1 KINRIRQVTSGEISIVAGA.....SSKDDVTITNLSASGAFVT 261

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame p2n.model -DEV=xlh
-O=/cn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST_QPWT-fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CGN 1 1 13135 @runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRBADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estrba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Score | Length | DB | ID | Description |
|------------|-------------|-------|--------|----|----------|-------------|
| 1 | 1384 | 100.0 | 6246 | 29 | AY413475 | Homo sapi |
| 2 | 1379 | 99.6 | 5970 | 29 | AY413476 | Pan trogl |
| 3 | 1370 | 99.0 | 3190 | 29 | AY413477 | Mus muscu |
| 4 | 951 | 68.7 | 728 | 14 | CB520016 | UI-M-GIO- |
| 5 | 943.5 | 68.2 | 5087 | 29 | AY405422 | Mus muscu |
| 6 | 925.5 | 66.9 | 5094 | 29 | AY405420 | Homo sapi |
| 7 | 865 | 62.5 | 759 | 13 | BU704105 | UI-M-F00- |
| 8 | 857 | 61.9 | 723 | 13 | BU704133 | UI-M-F00- |
| 9 | 847.5 | 61.2 | 737 | 14 | CB520473 | UI-M-F00- |
| 10 | 830.5 | 60.0 | 797 | 14 | CD349851 | UI-M-FY0- |
| 11 | 812.5 | 58.7 | 5069 | 29 | AY405421 | Pan trogl |
| 12 | 791 | 57.2 | 739 | 13 | BO769387 | UI-M-FY0- |
| 13 | 790 | 57.1 | 735 | 29 | AG045907 | Pan trogl |
| 14 | 720.5 | 52.1 | 764 | 14 | CB244782 | UI-M-FY0- |
| 15 | 704 | 50.9 | 588 | 28 | AZ625513 | IM0465104 |
| 16 | 684.5 | 49.5 | 648 | 13 | BU057564 | UI-M-F00- |
| 17 | 653.5 | 47.2 | 1494 | 28 | CC206192 | CH261-170 |
| 18 | 649 | 46.9 | 602 | 9 | AL654107 | AL654107 |
| 19 | 637 | 46.0 | 575 | 12 | BG800702 | 0052-06 M |
| 20 | 593 | 42.8 | 786 | 14 | CA752501 | UI-M-F00- |
| 21 | 575.5 | 41.6 | 551 | 12 | BI975935 | 484513 MA |
| 22 | 570 | 41.2 | 793 | 13 | BU205083 | 603104346 |
| 23 | 568 | 41.0 | 512 | 29 | CE005729 | tigr-gss- |
| 24 | 549 | 39.7 | 472 | 14 | CF531401 | UI-M-FY0- |
| 25 | 546 | 39.5 | 782 | 14 | CF537442 | UI-M-GIO- |
| 26 | 541.5 | 39.1 | 789 | 14 | CF537442 | UI-M-GIO- |
| 27 | 532.5 | 38.5 | 942 | 29 | CNS022KB | UI-M-FY0- |
| 28 | 521 | 37.6 | 774 | 29 | AG082821 | Tetraodon |
| 29 | 497 | 35.9 | 716 | 14 | CD803068 | Pan trogl |
| 30 | 496 | 35.8 | 794 | 14 | CB518529 | UI-M-GV0- |
| 31 | 492 | 35.5 | 854 | 14 | CF745232 | UI-M-GH0- |
| 32 | 443.5 | 32.0 | 730 | 9 | AU051196 | UI-M-GV0- |
| 33 | 355 | 25.7 | 933 | 29 | CNS041MM | AU051196 |
| 34 | 342 | 24.7 | 399 | 28 | AZ831924 | Tetraodon |
| 35 | 339.5 | 24.5 | 909 | 10 | BF980526 | 2M011N20 |
| 36 | 333 | 24.1 | 798 | 10 | BF144774 | 602304274 |
| 37 | 317.5 | 22.9 | 806 | 14 | CF539403 | BR144774 |
| 38 | 317 | 22.9 | 524 | 28 | AZ005582 | 601791454 |
| 39 | 317 | 22.9 | 601 | 13 | EX509644 | UI-M-GH0- |
| 40 | 316 | 22.8 | 1083 | 28 | CC239138 | RPCI-23-3 |
| 41 | 297 | 21.5 | 463 | 9 | AL928416 | AL928416 |
| 42 | 295 | 21.3 | 545 | 10 | BF075317 | AL928416 |
| 43 | 295 | 21.3 | 900 | 13 | BU267121 | BF075317 |
| 44 | 289 | 20.9 | 1071 | 29 | CNS059C0 | BU267121 |
| 45 | 286 | 20.7 | 602 | 9 | AL710647 | 603506284 |

ALIGNMENTS

RESULT 1
AY413475
LOCUS Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence, 6246 bp
DEFINITION AY413475 genomic survey sequence.
ACCESSION AY413475.1 GI:39769437
VERSION GSS.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 6246)

```

AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
             Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
             Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
             Adams,M.D. and Cargill,M.

TITLE        Inferring nonneutral evolution from human-chimp-mouse orthologous
             gene trios

JOURNAL      Science 302 (5652), 1960-1963 (2003)
PUBMED       14671302
REFERENCE    2 (bases 1 to 6246)
AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
             Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
             Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
             Adams,M.D. and Cargill,M.

TITLE        Direct Submission
JOURNAL      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
             Rockville, MD 20850, USA
COMMENT      This sequence was made by sequencing genomic exons and ordering
             them based on alignment.
FEATURES     Location/Qualifiers
             source          1..6246
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
             gene            1..6246
                        /locus_tag="HCM4903"

ORIGIN
Alignment Scores:
Pred. No.:      2,66-154      Length:      6246
Score:          1384.00      Matches:    261
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:     0
DB:             29           Gaps:         0

US-10-029-020-14_COPY_1490_1750 (1-261) x AY413475 (1-6246)

QY      1 LysilleAsnArgIleAArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
DB      2398 AAGATCAACCGGATGCAAGCTTAATACCCCATCTTCTTGCTGTGTGCTGATGGGAG 2457

QY      21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
DB      2458 CCCAGTGGCTGTGACTGTAATAATGATGCCAACTGTGATCTTTCTGGAGACGATGGT 2517

QY      41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
DB      2518 TATGCCAAGGATGCAAGCTTAATACCCCATCTTCTTGCTGTGTGCTGATGGGAG 2577

QY      61 LeuTyrValAlaAspLeuGlyAsnIleAArgIleAArgPheIleAArgLysAsnLysProPhe 80
DB      2578 CTCACGTCGGCGGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAGCCCTTC 2637

QY      81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrIleuPhe 100
DB      2638 CTCACACCCAGACATGATGAGCTGTCTTCCACCAATTGACCAAGAGCTCTATCTGTTT 2697

QY      101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrIleuTyrAsn 120
DB      2698 GATACCAACCGGCAAGCACTGTATACACCAAAAGCTTCCGCCACAGGAGACTACCTGTACAAAC 2757

QY      121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
DB      2758 TTCACCTACATCGGGAGCGGCACATCACTATCATCAGACACAAATGGCCACATGGTA 2817

QY      141 AsnValAArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
DB      2818 AATGTCGCCCGGAGACTCTACTGGGATGCCCTCTGGTGGTGGTCCAGATGGCCAGGTG 2877

QY      161 TyrTyrValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
DB      2878 TACTGGGTGACCATGGGCACCAACAGTGCACTCAAGAGTGTGCCACACAGGACACCGAG 2937

```

```

QY      181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
DB      2938 TTGGCCATGATGACATACCATGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAAC 2997

QY      201 GlyTyrThrThrPheTyrGluTyrAspSerPheGlyAArgLeuThrAsnValThrPhePro 220
DB      2998 GGATGGACAACATTTTATGAGTACGACAGCTTTGGCGCGCTTGACAAATGTGACCTTCCT 3057

QY      221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
DB      3058 ACTGGCCAGGTGACAGCTTTCCGAAGTGATACAGACAGTTCATGTCATGTCAGGTAGAG 3117

QY      241 ThrSerSerLysAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
DB      3118 ACCTCCAGCAAGGATGATGTACCAATACCAACCACTGTCTGCTCAGCGGCTTCTAC 3177

QY      261 Thr 261
DB      3178 ACA 3180

RESULT 2
LOCUS      AY413476                      5970 bp      DNA      linear      GSS 12-DEC-2003
DEFINITION Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
             genomic survey sequence.
ACCESSION  AY413476
VERSION    AY413476.1 GI:39769438
KEYWORDS   GSS.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
REFERENCE  1 (bases 1 to 5970)
            Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
            Science 302 (5652), 1960-1963 (2003)
            14671302
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Direct Submission
            Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
            This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES     Location/Qualifiers
             source          1..5970
                        /organism="Pan troglodytes"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9598"
                        <1..5970
                        /locus_tag="HCM4903"

ORIGIN
Alignment Scores:
Pred. No.:      9,67e-154      Length:      5970
Score:          1379.00      Matches:    260
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.62%      Mismatches: 0
Query Match:    99.64%      Indels:     0
DB:             29           Gaps:         0

US-10-029-020-14_COPY_1490_1750 (1-261) x AY413476 (1-5970)

QY      1 LysilleAsnArgIleAArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
DB      2122 AAGATCAACCGGATGCAAGCTTAATACCCCATCTTCTTGCTGTGTGCTGATGGGAG 2181

```

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
 Db 2182 CCACGTTGGCTGTGACTGTAAATATGATGCACTGTGATTTTCTTGAGACGATGGT 2241
 QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
 Db 2242 TATGCCAAGGATCAAAAGTTAAATATACCCCATCTTCCTGGCTGTGTCTGATGGGAG 2301
 QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
 Db 2302 CTCCTAGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACCAAGCCCTTC 2361
 QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
 Db 2362 CTCACACCCAGACATGTATGAGCTGTCTTACCAATTTACCAAGAGCTCTATCTGTTT 2421
 QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
 Db 2422 GATACCACTGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGCTACCTGTACAC 2481
 QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
 Db 2482 TTCACCTACCTACCTGGGACGGTGACATCACACTCATCACAGACAAATGGCAACATGGTA 2541
 QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
 Db 2542 AATGTCGCGCGAGACTCTACTGGGATGCCCCCTCTGGCTGGTCCAGATGGCCAGGTG 2601
 QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
 Db 2602 TACTGGGTGACCATGGGCACCAACAGTGACATCAAGAGTGTCACCAACAGGACACGAG 2661
 QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
 Db 2662 TTGGCCATGATGACATACATGGCAATTCGGCCCTTCTGGCAACCAAAAGCAGTGAATAAC 2721
 QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
 Db 2722 GAGTGACAAACATTTATGATGATACACAGCTTTGGCCGCTTGACAAATGTACCTTCCT 2781
 QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
 Db 2782 ACTGGCCAGGTGAGCAGTTTCGGAAGTGATACAGACAGTTCAGTGCATGTCAGGTAGAG 2841
 QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
 Db 2842 ACCTCCAGAGGATGTACCAATAACCAACCTGTCTGCTCGCTGAGTGCCTTCTAC 2901
 QY 261 Thr 261
 Db 2902 ACA 2904

RESULT 3
 AY413477
 LOCUS
 DEFINITION Mus musculus HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY413477
 VERSION AY413477.1 GI:39769439
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3190)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)

14671302
 2 (bases 1 to 3190)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment
 FEATURES
 Location/Qualifiers
 source
 1..3190
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>3190
 /locus_tag="HCM4903"
 gene
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.2e-153 Length: 3190
 Score: 1370.00 Matches: 258
 Percent Similarity: 99.62% Conservative: 2
 Best Local Similarity: 98.83% Mismatches: 1
 Query Match: 98.99% Indels: 0
 DB: 29 Gaps: 0
 US-10-029-020-14_COPY_1490_1750 (1-261) x AY413477 (1-3190)
 QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
 Db 2398 AAGATCAATCGCATCAGGCGGTCTACTAAGTGGTGAGATCTCACTGTTGCTGGTGCC 2457
 QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
 Db 2458 CCCAGTGGCTGTGACTGTAAATAATGATGCCAACTGTGACTCTCTCTGGAGATGATGTT 2517
 QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
 Db 2518 TACGCCAAGGATGCAAAAGCTCAATACCCATCGTCTTGGTGTGTGTGTCACGGGAG 2577
 QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
 Db 2578 CTCCTATGTGGCCGACCTGGGAAACATCCGAATTCGATTTATCCGGAAGAACCAAGCCCTTC 2637
 QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
 Db 2638 CTGAACACTCAGAACATGTACGAGCTATCTCCCCCATCGACACGAGGAGCTGTACTCTTT 2697
 QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
 Db 2698 GATACCAAGTGGCAAGCATCTGTACACTCAGAGCCCTACCCACAGGGGACTACTGTACAC 2757
 QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
 Db 2758 TTCCTTTACAGGGGACGGGACATCACATATCACCGCAACAATATGGCAACATGGTG 2817
 QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
 Db 2818 AACGTCGCCGAGACTCTACCGGATGCTCTCTGGCTGGTAGTCCAGATGGCCAGGTA 2877
 QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
 Db 2878 TACTGGGTAACTATGGCACCAACAGCGCATCAGAGTGTGCCACACAGGACACGAG 2937
 QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
 Db 2938 CTAGCCATGATGACCTACCATGGCAACTCTGGGCTCTCTGGCAACCAAAAGCAATGAAAC 2997
 QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
 Db 2998 GGGTGGCAACAGTTTATGATGATGACAGTTTGGTGGCTGCTGACAAACGTCACCTTTCCA 3057

| | | | |
|--|------------|---|------|
| QY | 221 | ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu | 240 |
| Db | 3058 | ACTGCCAGGTGAGCAGTTTCGCAAGCATACAGACAGCTCAGTCAGTCAGGAGTAGAG | 3117 |
| QY | 241 | ThrSerSerLysAspValThrLeThrThrAsnLeuSerAlaSerGlyAlaPheTyr | 260 |
| Db | 3118 | ACCTCAAGCAAGATGACGTCCACCATACCAACCTCTGCTTCGGGTGCCTCTAC | 3177 |
| QY | 261 | Tyr 261 | |
| Db | 3178 | ACC 3180 | |
| RESULT 4 | | | |
| CB520016 | | 728 bp mRNA linear EST 09-JUL-2003 | |
| LOCUS | | UI-M-GIO-cek-i-10-0-UI.r1 NIH BMAP_GIO Mus musculus cDNA clone | |
| DEFINITION | | IMAGE:6840755 5', mRNA sequence. | |
| ACCESSION | | CB520016 | |
| VERSION | | CB520016.1 GI:29353371 | |
| KEYWORDS | | EST. | |
| SOURCE | | Mus musculus (house mouse) | |
| ORGANISM | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | |
| REFERENCE | | 1 (bases 1 to 728) | |
| AUTHORS | | NIH-MGC http://mgc.nci.nih.gov/ . | |
| TITLE | | National Institutes of Health, Mammalian Gene Collection (MGC) | |
| JOURNAL | | Unpublished (1999) | |
| COMMENT | | Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html This clone was contributed by the Brain Molecular Anatomy Project (BMAP) | |
| FEATURES | | Seq primer: pYX-5. Location/Qualifiers 1..728 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6" /db_xref="taxon:10090" /clone="IMAGE:6840755" /tissue_type="whole brain" /dev_stage="embryo 13.5,14.5,16.5,17.5dpc" /lab_host="DH10B (T1 phage resistant)" /clone_lib="NIH BMAP_GIO" /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator." | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 1..13e-103 | Length: | 728 |
| Score: | 951.00 | Matches: | 179 |
| Percent Similarity: | 99.45% | Conservative: | 2 |
| Best Local Similarity: | 98.35% | Mismatches: | 1 |
| Query Match: | 68.71% | Indels: | 0 |
| DB: | 14 | Gaps: | 0 |
| US-10-029-020-14_COPY_1490_1750 (1-261) x CB520016 (1-728) | | | |
| QY | 80 | PheLeuAsnThrGlnAsnMetTyrGluLeuSerSerProileAspGlnGluLeuTyrLeu | 99 |
| Db | 7 | TTTCCTGAACACTCAGAACATGTACGAGCTATCTCCCCCATCGACCGAGAGCTGTACCTC | 66 |
| QY | 100 | PheAspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyr | 119 |
| Db | 67 | TTTGATACCAGTGGCAAGCATCTGTACACTCAGAGCTACCCACAGGGACTACCTGTAC | 126 |
| QY | 120 | AsnPheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMet | 139 |
| Db | 127 | AACTTCACCTTACACAGGGGACGGGACATCACACATATCACCCACACAAATGGCAACATG | 186 |
| QY | 140 | ValAsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGln | 159 |
| Db | 187 | GTGAACGTCCGGCGAGACTCTACCGGATGCCCTCTCTGGCTGGTAGTCCAGATGCCAG | 246 |
| QY | 160 | ValTyrTyrValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHis | 179 |
| Db | 247 | GTATATCGGTAAACCATGGGCACCAACAGCGCACTCAGAAAGTGTGACCAACAAAGGACAC | 306 |
| QY | 180 | GluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGlu | 199 |
| Db | 307 | GAGCTAGCCATGATGACCTACCATGGCACTCTGGCCTCTTGGCAACCAAAAGCAATGAA | 366 |
| QY | 200 | AsnGlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhe | 219 |
| Db | 367 | AAACGGTGGACACACGTTTATGATGATGACAGTTTGGTCGCTCGCTGACCAACAGTGACCTTT | 426 |
| QY | 220 | ProThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnVal | 239 |
| Db | 427 | CCAACTGGCCAGGTGAGCAGTTTCCGAAGCGATACAGACAGCTCAGTGCAGTGCAGGTA | 486 |
| QY | 240 | GluThrSerSerLysAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe | 259 |
| Db | 487 | GAGACCTCAAGCAAGATGACGTACCATTAACCAACCACTGTCTGCTCGGGTGCCTTC | 546 |
| QY | 260 | TyrThr 261 | |
| Db | 547 | TACACC 552 | |
| RESULT 5 | | | |
| AY405422 | | 5087 bp DNA linear GSS 12-DEC-2003 | |
| LOCUS | | Mus musculus HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence. | |
| DEFINITION | | AY405422 | |
| ACCESSION | | AY405422.1 GI:39761396 | |
| VERSION | | GSS. | |
| KEYWORDS | | Mus musculus (house mouse) | |
| SOURCE | | Mus musculus | |
| ORGANISM | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | |
| REFERENCE | | 1 (bases 1 to 5087) | |
| AUTHORS | | Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | |
| TITLE | | Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios | |
| JOURNAL | | Science 302 (5652), 1960-1963 (2003) | |
| PUBMED | | 14671302 | |
| REFERENCE | | 2 (bases 1 to 5087) | |
| AUTHORS | | Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | |
| TITLE | | Direct Submission | |
| JOURNAL | | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, | |

```
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
    source
        1..5087
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            <1..>5087
            /locus_tag="HCM2218"
ORIGIN
Alignment Scores:
Pred. No.:      2,2e-101      Length:      5087
Score:          943.50      Matches:      173
Percent Similarity: 79.39%      Conservative: 35
Best Local Similarity: 66.03%      Mismatches: 53
Query Match:      68.17%      Indels:      1
DB:              29          Gaps:      1
US-10-029-020-14_COPY_1490_1750 (1-261) x AY405422 (1-5087)
QY      1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
DB      1561 AAGATCAACCGCAATAGGAGGTCACGACAGCGGGAGATCTCTTAGTGGCTGGGATA 1620
QY      21 ProSerGlyCysAspCysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
DB      1621 CCTTCGGAAATGACTGCAAGAACGACGCACTGTGACTGTCTACCAAGCGGAGACGCGC 1680
QY      41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
DB      1681 TAGGCCAAAGATGCCAAACTCAATCGCGCTCTCTCCGCGCCCTCGCCAGATGGCACT 1740
QY      61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
DB      1741 CTGTACATTGTCAGATCTGGGAATATCAGGATCCGGGCGGCTTCGAAGAATAAACCTTTA 1800
QY      81 LeuAsnThrGlnAsnMetTyrCluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
DB      1801 CTGAACCTCAATGAACTTTTACGAAGTTGCTCTCCAACTGATCAAGAGCTCTACATCTTT 1860
QY      101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
DB      1861 GACATCAACGGTACTCACCAGTACACCGTGCAGCTGGTGGTCAAGGAGTACTTATATATAT 1920
QY      121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
DB      1921 TTTAGTTACAGCAATGACAAATGACATGACGTCACCGCTGTAAGTACAGCAATGGCAACCCCTC 1980
QY      141 AsnValArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
DB      1981 CGAATCCGAAGGATCCGAATCGGATCGGTCGGTGGTGGTCTCTCGATAACCAAGTG 2040
QY      161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
DB      2041 ATATGGTTGACCAATAGGACCAACCGGGTGTCTGAAACGATCAGCGCTCAGGGCTGGAA 2100
QY      181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
DB      2101 CTGGTTTGTGTACTTACCACTGGCAACAGTGGGCTTTTAGCCACCAAAAGTGACAAACT 2160
QY      201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
DB      2161 GGATGGACAACTTTTGTGATGATGACAGTGAAGTCCGCTCGGCAAGTGTACCTTCCCTC 2220
QY      221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
DB      2221 ACTGGGGTGGTTACAACTCCACGGGACATGACACAGGCTATCAGCTGGACATCGAG 2280
QY      241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
DB      2281 TCATCCAGCAGAGGAAGATGTGAGATCATCTTCGAACCTTGTCTCCATCGATTCCTTC 2340
```

```
QY      260 TyrThr 261
DB      2341 TACACC 2346

RESULT 6
AY405420
LOCUS
DEFINITION
    Homo sapiens HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
    genomic survey sequence.
ACCESSION
    AY405420
VERSION
    AY405420.1 GI:39761394
KEYWORDS
    GSS.
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 5094)
    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
    Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
    Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
    Adams,M.D. and Cargill,M.
    Inferring nonneutral evolution from human-chimp-mouse orthologous
    gene trios
    Science 302 (5652), 1960-1963 (2003)
JOURNAL
    PUBLISHED
    14671302
REFERENCE
    2 (bases 1 to 5094)
    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
    Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
    Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
    Adams,M.D. and Cargill,M.
    Direct Submission
    Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
    Rockville, MD 20850, USA
    This sequence was made by sequencing genomic exons and ordering
    them based on alignment.
FEATURES
    source
        1..5094
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            <1..>5094
            /locus_tag="HCM2218"
    gene
    ORIGIN

Alignment Scores:
Pred. No.:      3,25e-99      Length:      5094
Score:          925.50      Matches:      172
Percent Similarity: 78.24%      Conservative: 33
Best Local Similarity: 65.65%      Mismatches: 56
Query Match:      66.87%      Indels:      1
DB:              29          Gaps:      1
US-10-029-020-14_COPY_1490_1750 (1-261) x AY405420 (1-5094)
QY      1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
DB      1568 AAAATTAAACCGGATAAGGACAGTGCACACAGATGGAGAAATCTCTTAGTGGCGGAATA 1627
QY      21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
DB      1628 CCTTCAGAGTGTGACTGCAAAATGATGCCAACTGTGACTGTTCAGAGATGGAGATGGC 1687
QY      41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
DB      1688 TAGCCCAAGGATGCCAAACTCAGTGGCCCCCATCTCTCTGGCTGCTTCGAGATGGTACA 1747
QY      61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
DB      1748 CTGTATATTGAGATCTAGGGAATATCCGATCCGGCTGTGTCAAAGAATAACCCCTTTA 1807
QY      81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
DB      1807
```

```
Db 1808 CTTAACTCTATGAACCTTCTATGAAGTTGGCTCTCCAACTGATCAAGAACTCTACATCTTT 1867
QY 101 AspThrThrGlyHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 1868 GACATCAATGGTACTCACCATAATATCTGTAAGTTTGTAGTCACTGGTGAATACCTTTTCAAT 1927
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 1928 TTATGCTACAGCAGCATGATGATATATCTGCTGACAGACAGCATGGCAACACCCCT 1987
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProaspGlyGlnVal 160
Db 1988 AGAATTAGACGGGACCAACCAATCGCATGCCAGTTCGAGTGGTGTCTCTCTATACCAAGTG 2047
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 2048 ATATGGTTGACATAGCAACAAATGGATGTTTGAAAGCATGACTGCTCAAGGACTGGAA 2107
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 2108 TTAGTTTGTGTACTTACCATGCGAATAGTGGCTTTTACCCACTAAAGTGAAGAACT 2167
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 2168 GGATGGACAACGTTNNNNNCTATGACAGTCAAGTGTCTGTGACAAATGTTACGTTTCCA 2227
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 2228 ACTGGAGTGTGCACAACTCGATGGGACATGACCAAGCTATACAGTGGACATTCAG 2287
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 2288 TCATCTAGCCGAGNAGAGATGTGACATCACITCAATCTGTCTGCTCGATCTCTTC 2347
QY 260 TyrThr 261
Db 2348 TACACC 2353

RESULT 7
BU704105 759 bp mRNA linear EST 15-JUL-2003
LOCUS UI-M-FOO-bzr-d-20-0-UI.r1 NIH BMAP_F00 Mus musculus cDNA clone
DEFINITION IMAGE:6406243 5', mRNA sequence.
ACCESSION BU704105
VERSION BU704105.1 GI:23631890
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 759)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rcapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library prepared by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..759
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6406243"
```

```
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_F00"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TGAGAGACC. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
Program coordinator."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,77e-93 Length: 759
Score: 865.00 Matches: 161
Percent Similarity: 78.51% Conservative: 29
Best Local Similarity: 66.53% Mismatches: 52
Query Match: 62.50% Indels: 1
DB: 13 Gaps: 0
```

US-10-029-020-14_COPY_1490_1750 (1-261) x BU704105 (1-759)

```
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 35 AAGATCAACCGAATAAGCGCAGGTCAACAGACAGCGGGAGATCTCTTAGTGGCTGGGATA 94
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 95 CCTTCGGAATGTGACTGCAAGACGACGCGCACTGTGACTGTCTACCAAGCGGAGCGGC 154
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaGlyGlu 60
Db 155 TAGCCCAAGATGCCAAACTCAATGCGCGCTCTCTCTGCGCGCTCGCCAGATGGCACT 214
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 215 CTGTACATTGCAGATCTGGGAAATATCAGGATCGGGCGCTTCGAGAAATAAACCTTTA 274
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 275 CTGAACCTCAATGAACCTTTTACGAAGTTGCTCTCCAACTGATCAAGAGCTCTACATCTTT 334
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 335 GACATCAAGCGGTACTCACCAGTACACCGTGTGCTGACGGGTGTCTGAGAAATACCTATAAT 394
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 395 TTTTGTACAGCAATGATGATGACGTCAACCGCTGTACTGACAGCAATGGCAACACCTC 454
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProaspGlyGlnVal 160
Db 455 CGAATCGAAGGATCCGAATCGGATCGCGTGGGTGGTGTCTCTCTGATAACCAAGTG 514
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 515 ATATGGTTGACATAGGCACCAACCGGTGTCTGAAAGCATGACCGCTCAGCGCTGGAA 574
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 575 CTGGTTTGTGTACTTACCATGGCAACAGTGGCTTTTAGCCACCANAGTACCAACT 634
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
```


Db 635 GGATGACAAACATTTTTCATGATGACAGTGAAGGTGGCTGACGAATGTTACCTTCCCC 694
 QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
 Db 695 ACTGGGGTGGTTACAAACTG-CACGGGACATGGCAAGGCTATCAGGTGACATCGAG 753
 QY 241 ThrSer 242
 Db 754 TCATCC 759
 RESULT 8
 BU704133 723 bp mRNA linear EST 15-JUL-2003
 LOCUS UI-M-F00-bzr-j-10-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
 DEFINITION IMAGE:6406377 5', mRNA sequence.
 BU704133
 ACCESSION BU704133.1 GI:23631946
 VERSION BU704133
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 723)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..723
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6406377"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T1 phage clone resistant)"
 /clone_lib="NIH BMAP F00"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is TGAGAGAGCC. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN
 Alignment Scores:
 Pred. No.: 2,36e-92 Length: 723
 Score: 857.00 Matches: 156
 Percent Similarity: 80.09% Conservative: 25
 Best Local Similarity: 69.03% Mismatches: 45
 Query Match: 61.92% Indels: 0
 DB: 13 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x BU704133 (1-723)
 QY 1 LysIleAsnArgIleAArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
 Db 34 AAGATCAACCGAATGAAGCAGGTACAGACAGACGGGGAGATCTCTTAGTGGCTGGATA 93
 QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
 Db 94 CCTTCGGAATGTGACTGCAAGAACGACGCCAACTGTGACTGTACCAAGCGGAGACGCG 153
 QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
 Db 154 TACGCCAAGATGCCAAACTCAATCGCGCTCTCCCTGGCGCTCGCCAGATGGCACT 213
 QY 61 LeuTyrValAlaAspLeuGlyAsnIleAArgPheIleAArgLysAsnLysProPhe 80
 Db 214 CTGTACATTGCAGATCTGGGAATATCAGGATCCGGCGGTTTCGAAGATTAACCTTTA 273
 QY 81 LeuAsnThrGlnAsnMetTyrGlnLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
 Db 274 CTGAACCTCAATGAATTTTACGAGTTGCCTCTCCCACTGATCAAGAGCTCTACATCTTT 333
 QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
 Db 334 GACATCAACGGTACTCACCATGACCGTACCGTGGCTGGTCCAGGGTGACTACTATATAAT 393
 QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
 Db 394 TTTAGTTACAGCAATGACATGACAGTCACTGTAACCTGACAGCAATGGCAACACCTTC 453
 QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
 Db 454 CGAATCCGAGGATCCGAATCGGATCGCGTGGGTGGTGTCTCTCTGATAAACAGGTG 513
 QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
 Db 514 ATATGGTTGACCATYANGACCAACGGGTGTCTGAAAACATGACCGCTCAGGGCTGGAA 573
 QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
 Db 574 CTGGTTTGTGTTTACTTACCATGGCAACAGTGGCTTTTAGCCACCAAAAGTGACAAACT 633
 QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
 Db 634 GGATGGACACATTTTGTGACTATGACATGACAGTGAAGTGCCTGACGAGATGTTACCTTCCC 693
 QY 221 ThrGlyGlnValSerSer 226
 Db 594 ACTGGGTGGTTACACACC 711
 RESULT 9
 CB520473 737 bp mRNA linear EST 09-JUL-2003
 LOCUS UI-M-G10-cej-m-11-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
 DEFINITION IMAGE:6840468 5', mRNA sequence.
 CB520473
 ACCESSION CB520473.1 GI:29353828
 VERSION CB520473
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 737)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.

FEATURES
source
Location/Qualifiers
1. .737
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6840468"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH BMAP G10"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 3,39e-91 Length: 737
Score: 847.50 Matches: 155
Percent Similarity: 78.19% Conservative: 35
Best Local Similarity: 63.79% Mismatches: 51
Query Match: 61.24% Indels: 2
DB: 14 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x CB520473 (1-737)
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCyseAspCysPheSerGlyAspAspGly 40
DB 6 CCTTCGGAATGTACTGCAAGACGCGCCCACTGTGACTGTACCAACGCGAGACGGC 65
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
DB 66 TAGCCCAAGATGCCAACTCAATGCGCGCTCTCCCTGCGCGCTCGCCAGATGGCACT 125
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
DB 126 CTGTACATTGCAGATCTGGGAAATATACAGATCCGCGCGCTGTGCAAGAGATAAACCTTTA 185
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
DB 186 CTGAATCAATGAACTTTTACGAAGTTGCTCTCCCACTGATCAGAGCTCTACATCTTT 245
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
DB 246 GACATCAACGGTACTCACCAGTACACCGTGCAGCTGCTGCGGTGACTACCTATATAAT 305
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
DB 306 TTTAGTTACGAATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 365
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
DB 366 CGAATCCGAGGATCCGAATCGATCGCGGTGCGGTGCTCTCTGATTAACAGGTG 425
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180

Db 426 ATATGTTGACCATGACCAACCGGTGTCTCTGAAAGCATGACCGCTCAGGCGCTGGAA 485
QY 181 IeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuAlaThrLysSerAsnGluAsn 200
Db 486 CTGGTTTGTGTACTTACCATGGCAACAGTGGGCTTTTAGCCACCAAAAGTACGAACT 545
QY 201 GlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 546 GGATGGACAACTTTTGTGACTATGACAGTGAAGTGGCTGCGCTGACGAATGTACCTTCCC 605
QY 221 ThrGlyGln-ValSerSerPheArgSerAspThrAspSerSerValHisValGlnValG 240
Db 606 ACTGGGNGTGTGTACAACTGCAACCGGACATGCAAGGCTATCAGCGTGACATCGA 665
QY 240 uThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPh 259
Db 666 GTATCCACGACAGAGAGAGATGTCTAGCATCATCTTCTGAACTTGTCTCTCCATCGATTCCT 725
QY 259 eTyrThr 261
Db 726 CTACACC 732

RESULT 10
CD349851
LOCUS
DEFINITION
UI-M-FY0-cfr-n-12-0-UI.r1 NIH BMAP_FY0 Mus musculus CDNA clone
IMAGC:6855085 5', mRNA sequence.
CD349851
ACCESSION
CD349851.1 GI:31141366
VERSION
EST.
KEYWORDS
SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 797)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1. .737
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6855085"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH BMAP_FY0"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National

FEATURES
source
Location/Qualifiers
1. .737
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6855085"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH BMAP_FY0"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 4,31e-89 Length: 797
Score: 830.50 Matches: 154
Percent Similarity: 77.20% Conservative: 39
Best Local Similarity: 61.60% Mismatches: 56
Query Match: 60.01% Indels: 1
DB: 14 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x CD349851 (1-797)

QY 1 LysileAsnArgIleArgGlnValThrThrSerGlyGluileSerLeuValalaglyAla 20
DB 43 AAGATCAACCGCTAGCCAGTACCAACCAATGAGAGATCTGCTCTTACCGCGGGG 102
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
DB 103 GCCTCAGACTGTGACTGCAAAACCGATGCAACTGCATCTGCTACTCGGAGATGACGCT 162
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
DB 163 TACGCCACGCGCGCATCTAACTCGCGCTCTCTTACCGCTGCTCCAGATGGCACC 222
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
DB 223 ATCTACATTGACAGCTTGGGATATCCGGATCAGGCGGTGAGCAAAATAAACCCTGT 282
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
DB 283 CTTAACGCATTCAACACGATGAGGCTGCATCTCCGGGACACAGAAATGTACGTGTC 342
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
DB 343 AACCGTGTATGATCCATCAGTACACTGTGAGTCTGCTGAGTGGGAGTACTTGTACAA 402
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
DB 403 TTCATACACGGCTGCAATGACGTACCGAGTGTGATTGACAAACCGGAATTCCTCA 462
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
DB 463 AAGATCCGCGGGACAGCAGTGGATGCCCCGCCACCTGCTCATGCTGATCAATCAGATT 522
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
DB 523 ATCACCTTACCGTGGGACCAATGGAGCGCTCAAGCGGTGCTCCACTCAGAACCTGGAG 582
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
DB 583 CTGGGCTCATGATTATGATGGGAACACTGGATCTCTAGCCACCAAGAGTGTGAAACT 642
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
DB 643 GGATGGCAACTTTTATGACTATGACCAACCGAGGCGCTCTGNACATGTACCGCGCCC 702
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
DB 703 ACGGGCTAGTACGAGTCTGCACCGGGAATGGAGAAATCTATCACCATTGACATTGAG 762
QY 241 ThrSerSerLys---AspAspValThrIle 249
DB 763 AACTCCAACCGGATGATGACGTCACTGTG 792

RESULT 11

AY405421

LOCUS

DEFINITION Pan troglodytes HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION

AY405421

VERSION AY405421.1 GI:39761395

KEYWORDS

SOURCE Pan troglodytes (chimpanzee)

ORGANISM

REFERENCE 1 (bases 1 to 5069)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE

JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5069)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE

JOURNAL Direct Submission
COMMENT Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES

source

1..5069

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

<1..5069

/locus_tag="HCM2218"

gene

ORIGIN

Alignment Scores:

Pred. No.: 1.33e-85 Length: 5069

Score: 812.50 Matches: 155

Percent Similarity: 70.99% Conservative: 31

Best Local Similarity: 59.16% Mismatches: 75

Query Match: 58.71% Indels: 1

DB: 29 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x AY405421 (1-5069)

QY 1 LysileAsnArgIleArgGlnValThrThrSerGlyGluileSerLeuValalaglyAla 20

DB 1568 AAAATTAAACCGGATAAGGCGAGTCAACACAGATGAGAAATCTCTTAGTCCCGGAATA 1627

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40

DB 1628 CCTCAGAGTGTGACTGCAAAATGATGCCAACTGTGACTGTCCAGAGTGGAGATGGC 1687

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60

DB 1688 TACGCCAAGGATGCCAAACTCAGTGGCCCCCATCTCTCTGGCTGCTTCTCCAGATGGTACA 1747

QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80

DB 1748 CTGTATATTGACATCTAGGGAATATCCGATCCGGCTGTGTCAAGAAATAAGCCCTTTA 1807

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100

DB 1808 CTTNACTCTATGAACTTCTATGAAGTTGCGTCTCAACTGATCAAGAACTCTACATCTTT 1867

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120

DB 1868 GACATCAATGGTACTCACCATAATATCTGTAAAGTTTAGTCACTGGTGTATACCTTTACAT 1927

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140

DB 1928 TTTAGCTACAGCAATGACAAATGATATTACTGTGTGACAGACAGCAATGGCAACCCCTT 1987

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160

| | | | |
|--|--|---|-----------------|
| Db | 1988 | AGAAATTAGACGGGACCAACCAATCGCATGCCAGTTCGAGTGGTGTCTCCGATACCAACAGTG | 2047 |
| Qy | 161 | TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu | 180 |
| Db | 2048 | ATATGTTGACATAGGAACAATGATGTTTGAAAGCATGACTGCTCAAGGACTGGAA | 2107 |
| Qy | 181 | LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn | 200 |
| Db | 2108 | TTAGTTTNGCTTACTTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN | 2167 |
| Qy | 201 | GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro | 220 |
| Db | 2168 | NN | 2227 |
| Qy | 221 | ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu | 240 |
| Db | 2228 | ACTGGAGTGGTCAACAACCTGCATGGGACATGGCAAGGCTATCATCAGTGGCAATGAG | 2287 |
| Qy | 241 | ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe | 259 |
| Db | 2288 | TCAATCAGCGGAGGAGGAAGATGTCAGATCATTCAAATCTGCTCGATGCTTTTC | 2347 |
| Qy | 260 | TyrThr | 261 |
| Db | 2348 | TACACC | 2353 |
| RESULT 12 | | | |
| BO769387 | | | |
| LOCUS | 739 bp | mRNA | linear |
| DEFINITION | UT-M-F10-bvg-i-13-01.r1 NIH BMAP_F10 Mus musculus cDNA clone | | EST 26-JUL-2002 |
| ACCESSION | BO769387 | | |
| VERSION | BO769387.1 | GI:21977861 | |
| KEYWORDS | EST. | | |
| SOURCE | Mus musculus (house mouse) | | |
| ORGANISM | Mus musculus | | |
| REFERENCE | | | |
| AUTHORS | Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| TITLE | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| JOURNAL | 1. (bases 1 to 739) | | |
| COMMENT | NH-MGC http://mgi.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs@nsl.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov This clone was contributed by the Brain Molecular Anatomy Project (BMAP) | | |
| FEATURES | | | |
| source | | | |
| Location/Qualifiers | | | |
| 1..739 | | | |
| /organism="Mus musculus" | | | |
| /mol_type="mRNA" | | | |
| /strain="C57BL/6" | | | |
| /db_xref="taxon:10090" | | | |
| /clone="IMAGE:6400980" | | | |
| /tissue_type="whole brain" | | | |
| /lab_stages="embryo 12.5dpc" | | | |
| /lab_hosts="DH10B (T1 phage resistant)" | | | |
| /clone_lib="NIH BMAP_F10" | | | |
| /note="Organ: Brain; Vector: pYX-Asc; Site: 1; EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned | | | |

| | | | |
|--|--|--|-----------------|
| directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator." | | | |
| ORIGIN | | | |
| Alignment Scores: | 2.19e-84 | Length: | 739 |
| Pred. No.: | 791.00 | Matches: | 149 |
| Score: | 78.44% | Conservative: | 22 |
| Percent Similarity: | 68.35% | Mismatches: | 47 |
| Best Local Similarity: | 57.15% | Indels: | 1 |
| Query Match: | 13 | Gaps: | 0 |
| DB: | | | |
| US-10-029-020-14_COPY_1490_1750 (1-261) x BO769387 (1-739) | | | |
| Qy | 1 | LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla | 20 |
| Db | 86 | AGATCAACCGATAGGCGAGGTACACAGACGGGGAGATCTCTTAGTGGCTGGATA | 145 |
| Qy | 21 | ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly | 40 |
| Db | 146 | CCTTCGGAATGTGACTGCAAGAACGACGCCNACTGTGACTGTACCAAAACGGAGACGC | 205 |
| Qy | 41 | TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu | 60 |
| Db | 206 | TAGCCCAAGATGCCAAACTCAATGGCCGCTCTCTCGCCGCCCTCCAGATGGCACT | 265 |
| Qy | 61 | LeuTyrValAlaAspLeuGlyAsnIleArgPheIleArgLysAsnLysProPhe | 80 |
| Db | 266 | CTGTACATTGCAGATCTGGGAATATCAGGATCCGGCCGCTTTCGAAGAATAAACCTTA | 325 |
| Qy | 81 | LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe | 100 |
| Db | 326 | CTGAACCTCAATGAACCTTTTACGAAGTTGCTCTCCAACTGATCAAGAGCTCTACATCTT | 385 |
| Qy | 101 | AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn | 120 |
| Db | 386 | GATCAACCGTACTACAGTACACCGTACGCTGCTGCTCGGGTGTACTTACCTATATAT | 445 |
| Qy | 121 | PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal | 140 |
| Db | 446 | TTTAGTTACAGCAATGACATGACGTCACCGCTGTAACGACCAATGGCAACACCGCTC | 505 |
| Qy | 141 | AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal | 160 |
| Db | 506 | CGAATCCGAAGGATCCGAATCCGATCCGCGTGGGTGTCTCTCTGATAACAGGTG | 565 |
| Qy | 161 | TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu | 180 |
| Db | 566 | ATATGGTTGACCATAGGACCAACCGGCTGTCTGAAGAGCATGACCGCTAGGGCTGGAA | 625 |
| Qy | 181 | LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn | 200 |
| Db | 626 | CTGGTTTNGTTTACTTACCATGGCAACAGTGGGCTTTTAGCCACCAAAAGTGACGAACT | 685 |
| Qy | 201 | GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThr | 218 |
| Db | 686 | GGATGGAC-ACATTTTTTGTACTATGACAGTGAGGTGCGCTGACGAATGTACT | 738 |
| RESULT 13 | | | |
| AG045907 | | | |
| LOCUS | 735 bp | DNA | linear |
| DEFINITION | Pan troglodytes DNA, clone: PTB-024P12.F, genomic survey sequence. | | GSS 02-NOV-2001 |
| ACCESSION | AG045907 | | |
| VERSION | AG045907.1 | GI:16582799 | |
| KEYWORDS | GSS. | | |
| SOURCE | Pan troglodytes (chimpanzee) | | |
| ORGANISM | Pan troglodytes | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB

2 (bases 1 to 735)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end clone was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. 735
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-024P12.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.: 2,86e-84 Length: 735
Score: 790.00 Matches: 147
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.32% Mismatches: 0
Query Match: 57.08% Indels: 0
DB: 29 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AG045907 (1-735)

QY 60 GluLeuTyrValAlaAspLeuGlyAsnLeuArgPheIleArgGlyAsnLysPro 79
|||||
Db 135 GAGCTCTACGTGGCGCGACCTTGGGAACATCGGAATTCGGTTTATCCGGAAGAACAGCCT 194

QY 80 PheLeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeu 99
|||||
Db 195 TTCTCTACACCCAGACATGATGAGCTGTCTTCCAAATGACCAAGAGCTCTATCTG 254

QY 100 PheAspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyr 119
|||||
Db 255 TTTTGATACCACTGGCAAGCCTGTACACCCAAAGCCTGCCACAGAGACTACTCTGAC 314

QY 120 AsnPheThrTyrThrGlyAspGlyAspIleThrLeulleThrAspAsnGlyAsnMet 139
|||||
Db 315 AACCTCACCTTACACTGGGGCGGTGACATCATCTCATCAGACAAACAATGGCAACATG 374

QY 140 ValAsnValArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyCln 159
|||||
Db 375 GTAATGTCCTCCAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAG 434

QY 160 ValTyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHis 179
|||||
Db 435 GTGTACTGGGTGACCATGGGCAACACAGTCACTCAAGAGTGTGACCAACACAGAGCAC 494

QY 180 GluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGlu 199
|||||
Db 495 GAGTTGCCCATGATGACATACCATGGCAATCTGGCCTCTCTGGCAACCAAAAGCAGTGAA 554

QY 200 AsnGlyTyrThrThrPheTyrGlu 207
|||||
Db 555 AACGGATGACCAACATTTTATGAG 578

RESULT 14
CB244782

LOCUS
CB244782

DEFINITION
UI-M-FY0-cdr-1-01-0-UI-r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:6833522 5', mRNA sequence.

ACCESSION
CB244782

VERSION
EST.

KEYWORDS
CB244782.1 GI:28366426

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus

REFERENCE
1 (bases 1 to 764)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1. 764
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6833522"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_FY0"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 7.2e-76 Length: 764
Score: 720.50 Matches: 146
Percent Similarity: 72.33% Conservative: 37
Best Local Similarity: 57.71% Mismatches: 56
Query Match: 52.06% Indels: 16
DB: 14 Gaps: 2

US-10-029-020-14_COPY_1490_1750 (1-261) x CB244782 (1-764)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
|||||
Db 36 AAGATCAACCGCTACGCCAAGTACCAATGGAGATCTGCTCTTAGCGGGGCG 95

| | | | |
|----|-----|---|-----|
| QY | 21 | ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly | 40 |
| Db | 96 | GCCTCAGACTGTGACTGCAAAAACGATGCACTGCATCTGCTACTCGGGAGATGACGCT | 155 |
| QY | 41 | TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu | 60 |
| Db | 156 | TACGGCAGCGGCATCCTAAACTCCGCGCTCTCTTAGCCGTGGCTCCAGATGGCACC | 215 |
| QY | 61 | LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe | 80 |
| Db | 216 | ATCTACATTGACAGCTTTGGGAATATCCGATCAGGCGGTGACCAAAATAAACCGTT | 275 |
| QY | 81 | LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe | 100 |
| Db | 276 | CTTAACGCATTCAACACGATGAGCTGCATCTCCGGAGAACACAGGAATTTGTACGTTC | 335 |
| QY | 101 | AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn | 120 |
| Db | 336 | AACGCTGATGGTATCCATCAGTACACTGTGAGTCTGTGACTGGGGAGTACTTGTCAT | 395 |
| QY | 121 | PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal | 140 |
| Db | 396 | TTTCACATACAGCGCTGACAAATGACGTCAACCGATTGATGACCAACACGGGAATTCCTA | 455 |
| QY | 141 | AsnValArgAspSerThrGlyMetProLeuThrLeuValValProAspGlyGlnVal | 160 |
| Db | 456 | AAGATCCGCGGGACAGCAGTGGCATGCCCGCCACCTGCTCATGCTGATAATCAGATT | 515 |
| QY | 161 | TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu | 180 |
| Db | 516 | ATCACCTTTACCGTGGGACCAATGGAGGCCTCAAGCCGTTCCATCTCAGACCTGGAG | 575 |
| QY | 181 | LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn | 200 |
| Db | 576 | CTGGGCCTCATGACTATGATGGGAACACTGGACTCTCTAGCCACCAAGAGTGTGAACT | 635 |
| QY | 201 | GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro | 220 |
| Db | 636 | GGATGGAC-ACTTTTTATGACTATGACACGGAGGGCGTCT---GACAATGTGACCCG-CCC | 690 |
| QY | 221 | ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu | 240 |
| Db | 691 | ACGGGGCTAGTGACC-----ATCTGC | 711 |
| QY | 241 | ThrSerSerLysAspAspValThrIleThrThrAsnLeu | 253 |
| Db | 712 | ACCGGAATGGGAATCTATCTACCATGATGACATTGAGATC | 750 |

| | |
|------------|--|
| RESULT | 15 |
| LOCUS | AZ625513/c |
| DEFINITION | Mus musculus genomic clone UGCLM0465I04 F, genomic survey sequence. |
| ACCESSION | AZ625513 |
| VERSION | AZ625513.1 |
| KEYWORDS | GSS. |
| SOURCE | Mus musculus (house mouse) |
| ORGANISM | Mus musculus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 598) Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.:Weiss,R. Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts Unpublished (2000) Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT, UT |
| REFERENCE | |
| AUTHORS | |
| TITLE | |
| JOURNAL | |
| COMMENT | |

```

84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0465 row: 1 column: 04
Seq primer: CGTTGTAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 588.
Location/Qualifiers
1. .588
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0465104"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid p1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

ORIGIN

| | | | |
|--|----------|---|-----|
| Alignment Scores: | 4.56e-74 | Length: | 588 |
| Pred. No.: | 704.00 | Matches: | 132 |
| Score: | 98.53% | Conservative: | 2 |
| Percent Similarity: | 97.06% | Mismatches: | 1 |
| Best Local Similarity: | 50.87% | Indels: | 1 |
| Query Match: | 28 | Gaps: | 0 |
| DB: | | | |
| US-10-029-020-14_COPY_1490_1750 (1-261) x AZ625513 (1-588) | | | |
| Qy | 73 | PheIleArgLysAsnLysProPheLeuAsnThrGlnAsnMetTyrGluLeuSerSer-Pr | 92 |
| Db | 586 | TTTATCCGGAAGAACAGCCTTTCTCTGAACACTCAGAACATGTGTACGAGCTATCTCCCCC | 527 |
| Qy | 92 | oileAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLeuTyrThrGlnSerLe | 112 |
| Db | 526 | CATGCACCAGAGGCTGTACCTCTTTTGATACCGTGGCAGGCACTCTGTACACTCAGAGCCT | 467 |
| Qy | 112 | uProThrGlyAspTyrLeuTyrAsnPheThrTyrThrGlyAspGlyAspIleThrLeuI | 132 |
| Db | 466 | ACCACAGGGGACTACCTGTGTACAACTTCACCTTACACAGGGACGGGACATCACACATAT | 407 |
| Qy | 132 | eThrAspAsnAsnGlyAsnMetValAsnValArgAspSerThrGlyMetProLeuTr | 152 |
| Db | 406 | CACCGACAAACATGGCAACATGGGTGAACGTCGGCGAGACTCTACCGGAGTGCCTCTCTG | 347 |
| Qy | 152 | pLeuValValProAspGlyGlnValTyrTrpValThrMetGlyThrAsnSerAlaLeuI | 172 |
| Db | 346 | GCTGGTAGTCCAGATGCCAGGTATCTGGGTACCACTGGGCACCAACAGCGCACTCAG | 287 |
| Qy | 172 | sServValThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLe | 192 |
| Db | 286 | AAGTGTGTGACCCACAAGGACACGAGCTAGCATGATGACCTACCATGGCACTCTGGCCT | 227 |

Qy 192 uLeuAlaThrLySerAsnGluAsnGlyTrpThrThrPheTyrGlu 207
|||
Db 226 CTGGCAACCAAGCAATGAAACGGGTGGACACGTTTATGAG 181

Search completed: August 14, 2004, 18:03:46
Job time : 3083.96 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 442.726 Seconds

(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_1490_1750

Perfect score: 1384

Sequence: 1 KINRIRQVTTSGEISLVAGA.....SSKDDVTITNLSASGAFVT 261

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/Cgn2_1/USPTO_spool/US10029020/runat_06082004_112215_29265/app query.fasta_1.3519
-DB=N Geneseq_29Jan04 -QEXT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pt0 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020.cgn_1_1868@runat_06082004_112215_29265 -NCFU=6 -ICFU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOB=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 1384 | 100.0 | 8354 | 6 | AB552100 Human TEN |
| 2 | 1384 | 100.0 | 8438 | 6 | ABN85378 Human NOV |
| 3 | 1380 | 99.7 | 8645 | 6 | AB578652 Human CDN |
| 4 | 935.5 | 67.6 | 8473 | 6 | ABQ82345 Human NOV |
| 5 | 935.5 | 67.6 | 8487 | 6 | ABQ82346 Human NOV |
| 6 | 935.5 | 67.6 | 8645 | 6 | ABQ82344 Human NOV |
| 7 | 935.5 | 67.6 | 8675 | 6 | ABQ82343 Human NOV |
| 8 | 898 | 64.9 | 12879 | 6 | ABK92230 Prostate |

| | | | | | | |
|----|-------|------|--------|---|-------------|---------------------|
| 9 | 898 | 64.9 | 13202 | 4 | AAK51828 | Aak51828 Human pol |
| 10 | 863.5 | 62.4 | 9058 | 7 | ACC72051 | Acc72051 BC0205A |
| 11 | 863.5 | 62.4 | 9695 | 7 | ACC72052 | Acc72052 BC0205B |
| 12 | 863.5 | 62.4 | 9729 | 5 | AAS14089 | Aas14089 Human FCT |
| 13 | 863.5 | 62.4 | 9729 | 5 | ADB32028 | Adb32028 Human FCT |
| 14 | 863.5 | 62.4 | 9826 | 5 | AAS14085 | Aas14085 Human FCT |
| 15 | 863.5 | 62.4 | 9826 | 5 | ADB32023 | Adb32023 Human FCT |
| 16 | 812.5 | 58.7 | 4245 | 7 | AAL60066 | Aal60066 Human PCO |
| 17 | 522.5 | 37.8 | 2434 | 5 | AAS71215 | Aas71215 DNA encod |
| 18 | 390.5 | 28.2 | 10242 | 4 | ABL29075 | Ab129075 Drosophil |
| 19 | 390.5 | 28.2 | 17131 | 4 | ABL29074 | Ab129074 Drosophil |
| 20 | 389 | 28.1 | 3310 | 6 | ABK34713 | Abk34713 Human CDN |
| 21 | 270.5 | 19.5 | 776 | 4 | AHH05747 | Aah05747 Human CDN |
| 22 | 270.5 | 19.5 | 3614 | 4 | AHH14183 | Aah14183 Human CDN |
| 23 | 193 | 13.9 | 3910 | 4 | ABLO4849 | Ab104849 Drosophil |
| 24 | 193 | 13.9 | 6318 | 4 | ABLO4848 | Ab104848 Drosophil |
| 25 | 131.5 | 9.5 | 548 | 6 | ABQ37165 | Abq37165 Oligonuc1 |
| 26 | 131.5 | 9.5 | 548 | 6 | ABQ37164 | Abq37164 Oligonuc1 |
| 27 | 119.5 | 8.6 | 4716 | 7 | ACA45361 | Aca45361 Prokaryot |
| 28 | 115 | 8.3 | 31960 | 8 | ACD18998 | Acad18998 E. coli 0 |
| 29 | 115 | 8.3 | 35910 | 9 | ADC00933 | Adc00933 Enterohae |
| 30 | 112.5 | 8.1 | 2415 | 8 | ADA30132 | Ada30132 DNA encod |
| 31 | 111.5 | 8.1 | 1235 | 6 | ABO68530 | Abq68530 Listeria |
| 32 | 110.5 | 8.0 | 5394 | 6 | ABQ71024 | Abq71024 Listeria |
| 33 | 110 | 7.9 | 6567 | 4 | AAS59636 | Aas59636 Propionib |
| 34 | 110 | 7.9 | 6567 | 7 | ACF64565 | Acf64565 Propionib |
| 35 | 109.5 | 7.9 | 1764 | 4 | ABLO4854 | Ab104855 Drosophil |
| 36 | 109.5 | 7.9 | 4925 | 4 | ABLO4854 | Ab104855 Drosophil |
| 37 | 109.5 | 7.9 | 6392 | 6 | ABQ67876 | Abq67876 Listeria |
| 38 | 109.5 | 7.9 | 6504 | 6 | ABQ69406 | Abq69406 Listeria |
| 39 | 109.5 | 7.9 | 110000 | 6 | ABQ69245_04 | Continuation (5 of |
| 40 | 109.5 | 7.9 | 110000 | 6 | ABQ67197_03 | Continuation (4 of |
| 41 | 108 | 7.8 | 4155 | 7 | ACA44188 | ACA44188 Prokaryot |
| 42 | 107.5 | 7.8 | 2532 | 7 | ACA52160 | ACA52160 Prokaryot |
| 43 | 107.5 | 7.8 | 4065 | 7 | ACA52159 | ACA52159 Prokaryot |
| 44 | 107 | 7.7 | 4497 | 7 | ACF72220 | ACF72220 Photorhab |
| 45 | 107 | 7.7 | 96798 | 7 | ACF65387_6 | Continuation (7 of |

ALIGNMENTS

RESULT 1

ABS52100
ID ABS52100 standard; DNA; 8354 BP.

XX XX ABS52100;

XX XX 05-NOV-2002 (first entry)

XX XX Human TEN-M4-like gene.

DE DE Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

XX XX Human TEN-M4-like gene.

OS Homo sapiens.
 XX Key Location/Qualifiers
 FT FH 4. .8395
 FT CDS /*tag= a
 FT /trans_except= (pos: 1138..1147,aa:Met)
 FT /product= "NOV1 protein"
 XX W0200255704-A2.
 XX 18-JUL-2002..
 XX 09-JAN-2002; 2002WO-US000554.
 XX 09-JAN-2001; 2001US-0260417P.
 XX 10-JAN-2001; 2001US-0260831P.
 XX 28-FEB-2001; 2001US-0272338P.
 XX 09-MAR-2001; 2001US-0274876P.
 XX 18-APR-2001; 2001US-0284704P.
 XX (CURA-) CURAGEN CORP.
 XX Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
 PI Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CM;
 PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
 PI Alsbrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
 PI Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
 PI Gunther E, Stone DJ;
 XX WPI; 2002-590674/53.
 DR P-PSDB; ABB98401.
 XX NOVX polypeptides and encoding polynucleotides, useful for preventing or
 PT treating NOVX-associated disorders e.g. cancer, inflammation, or
 PT Alzheimer's disease, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX Claim 9; Page 8-9; 358pp; English.
 XX The present sequence is a coding sequence for a NOV protein. The NOV
 CC proteins and coding sequences are useful for treating or preventing NOV-
 CC associated disorders or in the manufacture of a medicament for treating
 CC the disorders, such as cancer, heart disease, inflammation, autoimmune
 CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
 CC IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
 CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
 CC and other wasting disorders associated with chronic diseases. NOV1 is a
 CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
 XX SQ Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,91e-133 Length: 8438
 Score: 1384.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-029-020-14_COPY_1490_1750 (1-261) x ABB985378 (1-8438)
 QY 1 LyslleAsnArglleArgGlnValThrSerGlyGluLeuSerLeuValAlaGlyAla 20
 Db 4508 AAGATCAACCGCATCAGGAGGTCACCACTAGTGAGAGATCTCCTCGTGGGGCC 4567
 QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
 Db 4568 CCCAGTGGCTGTGACTGTAAATGATGCCCACTGTGATTTCTGGAGACGATGGT 4627
 QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
 Db 4628 TATGCCNAGGATGCAGAAATTAATACCCCATCTTCCTTGGCTGTGTGTGCTGGGAG 4687

QY 61 LeuTyrValAlaAspLeuGlyAsnLleArglleArgPheIleArgLysAsnLysProPhe 80
 Db 4688 CTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAGCCTTTC 4747
 QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGlnLeuTyrLeuPhe 100
 Db 4748 CTCACACCCAGAACATGATGAGCTGCTCTTCCACCAATTGACCCAGAGCTCTATCTGTTT 4807
 QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrThrGlyAspTyrLeuTyrAsn 120
 Db 4808 GATACCCCGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTTACCTGTAAAC 4867
 QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeulleThrAspAsnAsnGlyAsnMetVal 140
 Db 4868 TTCACCTACTGGGGCGGCGACATCATCTCATCACAGACACAAATGGCAACATGGTA 4927
 QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
 Db 4928 AATGTCGGCCGAGACTCTACTGGATGCCCTCTCGCTGGTGGTCCAGATGCCAGTGS 4987
 QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
 Db 4988 TACTGGGTGACCATGGGCACCAACAGTCAGCTCAAGAGTGTGACCAACAGGACACGAG 5047
 QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
 Db 5048 TTGGCCATGATGACATACATCCGCAATTCGGCCCTCTTGGCAACCAAAAGCAATGAAAC 5107
 QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
 Db 5108 GGATGGACAAATTTATGATGACAGCTTTGGCCCTGACAAATGTGACCTTCCT 5167
 QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
 Db 5168 ACTGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCAGTGCATGTCAGGTAGAG 5227
 QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
 Db 5228 ACCTCCAGCAAGGATGATGTCCATATACCACCACTGTCTGCTCGCCGCGCTTCTAC 5287
 QY 261 Thr 261
 Db 5288 ACA 5290
 RESULT 3
 ABS78652
 ID ABS78652 standard; cDNA; 8645 BP.
 XX AC ABS78652;
 XX DT 16-DEC-2002 (first entry)
 XX DE Human cDNA encoding CGDD10, INCYTE 7488573CB1.
 XX KW Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
 XX KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
 XX KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
 XX KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
 XX KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
 XX KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
 XX KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
 XX KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
 XX KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
 XX KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
 XX KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
 XX KW rheumatoid arthritis.
 XX OS Homo sapiens.
 XX WO200272830-A2.
 XX 19-SEP-2002.
 XX PD

PF 08-FEB-2002; 2002WO-US0037115.
XX 09-FEB-2001; 2001US-0268111P.
PR 23-FEB-2001; 2001US-0271175P.
PR 08-MAR-2001; 2001US-0274503P.
PR 09-MAR-2001; 2001US-0274552P.
XX (INCY) INCYTE GENOMICS INC.
PA Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
XX Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Dugan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Walla NK;
XX WPI; 2002-723356/78.
DR P-PSDB; ABG97359.
XX New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.
XX Claim 5; Page 175-178; 181pp; English.
XX The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death) a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide and
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotides,
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
XX a CGDD protein.
SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 1.59e-132 Length: 8645
Score: 1380.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.71% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x ABG78652 (1-8645)
QY 1 LysileAsnArgileArgGlnValThrThrSerGlyGluileSerLeuValAlaGlyAla 20
DB 4551 AAGATCAACCGCATCAGCAGGTCACCACTAGTCGAGAGATCTCACTGCTGGGGGCC 4610
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
DB 4611 CCCAGTGGCTGTGACTGTAAAAATGATGCCCACTGTGATTGTTTCTTGGAGACATGGT 4670

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
DB 4671 TATGCCAAGGATGCAAGTAAATACCCCATCTTCTTGGCTGTGTGTTGTTATGGGGAG 4730
QY 61 LeuTyrValAlaAspLeuGlyAsnileArgileArgPheIleArgLysAsnLysProPhe 80
DB 4731 CTCTACGTGGCGGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGACAAAGCCTTTC 4790
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGlnLeuTyrLeuPhe 100
DB 4791 CTCACACCCAGAACATGATGAGCTGTCTTCCACCAATGACAGGAGCTCTATCTGTTT 4850
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
DB 4851 GATACACCGGCAAGCACCTGTACACCCCAAGCCTGCCACAGGAGACTACCTGTACAAC 4910
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
DB 4911 TTCACTACACTGGGGACGCGGACATCACACTCATCAGACACACATGCGCAACATGGTA 4970
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
DB 4971 ATGTCCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTCCACATGCCAGGTG 5030
QY 161 TyrTyrValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
DB 5031 TACTGGGTACCATGGGCACCAACAGTGCACCTCAAGAGTGTGACCAACAAGACACGAG 5090
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
DB 5091 TTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAAC 5150
QY 201 GlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
DB 5151 GGTATGACCAACATTTTATGAGTACGACAGCTTTGGCGCGCTGACAAATGTGACTTCCCT 5210
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
DB 5211 ACTGGCCAGGTGAGCAGTTCCTCGAAGTGATACAGACAGTTTCAGTGCATGTCAGGTAGAG 5270
QY 241 ThrSerSerLysAspAspValThrThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
DB 5271 ACCTCCAGCAAGATGATGTTCACCAATAACCAACCACTGTCTGCTCAGGCGCTTCTAC 5330
QY 261 Thr 261
DB 5331 ACA 5333

RESULT 4
ABQ82345
ID ABQ82345 standard; cDNA; 8473 BP.
XX AC ABQ82345;
XX 17-DEC-2002 (first entry)
DT Human NOV15c encoding cDNA SEQ ID NO:39.
DE Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; vitucide;
KW antianaemic; antibacterial; protozoicide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX Homo sapiens.
OS
XX Location/Qualifiers
FH

KW Human; NOXV; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FH 299 .8140
FT /*tag= a
FT /product= "NOV15d"
XX

XX W0200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049576.

XX 29-DEC-2000; 2000US-0258928P.
XX 02-JAN-2001; 2001US-0259415P.
XX 04-JAN-2001; 2001US-0259785P.
XX 20-FEB-2001; 2001US-0269814P.
XX 09-MAR-2001; 2001US-0279863P.
XX 29-MAR-2001; 2001US-0279832P.
XX 29-MAR-2001; 2001US-0279833P.
XX 13-APR-2001; 2001US-0283889P.
XX 18-APR-2001; 2001US-0284477P.
XX 25-APR-2001; 2001US-0286683P.
XX 29-MAY-2001; 2001US-0294080P.
XX 16-AUG-2001; 2001US-0312915P.
XX 17-AUG-2001; 2001US-0313325P.
XX 17-SEP-2001; 2001US-0322699P.
XX 26-NOV-2001; 2001US-0333350P.

XX (CURA-) CURAGEN CORP.

XX Spytek KA, Li L, Wolenc AR, Vernet CM, Eisen A, Liu X;
XX Malyankar U, Shinkets RA, Tchernev VA, Spaderna SK, Gorman L;
XX Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
XX Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
XX Gunther E, Smithson G, Millet I, Macdougall JR;

XX WPI; 2002-732706/79.
XX P-PSDB; ABP53589.

XX New NOXV polypeptides and polynucleotides useful for treating NOXV-
XX associated disorders, such as cancers, neurological disorders, disorders
XX of vesicular transport, gastrointestinal disorders, and autoimmune
XX diseases.

XX Claim 8; Page 123-125; 444pp; English.

XX The present invention describes novel human proteins designated NOXV,
XX where X is 1 to 20 e.g. NOXV1. NOXV sequences can have neuroprotective,
XX cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
XX antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
XX immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
XX protozoacide and antihelminthic activities, and can be used in gene
XX therapy. The NOXV proteins, nucleotides or antibodies can be used in the
XX manufacture of a medicament for treating a syndrome associated with a
XX human disease selected from NOXV-associated disorder, such as cancers
XX (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
XX ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
XX ischaemic cerebrovascular disease, Alzheimer's disease or pick's
XX disease), disorders of vesicular transport (e.g. cystic fibrosis,
XX diabetes mellitus, Grave's disease, or goitre), gastrointestinal

CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOXV proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOXV nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15d, which is
CC located on chromosome 4
XX
XX Sequence 8487 BP; 2306 A; 2029 C; 2148 G; 2001 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 2.34e-86 Length: 8487
Score: 935.50 Matches: 172
Percent Similarity: 79.01% Conservative: 35
Best Local Similarity: 65.65% Mismatches: 54
Query Match: 67.59% Indels: 1
DB: Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x ABQ82346 (1-8487)

QY 1 LysileAsnArgileArgGlnValThrThrSerGlyGluileSerLeuValAlaGlyAla 20
Db 4313 AAAATTAACCGGATAAGCGCAGTCACACAGATGGAGAAATCTCTTAGTGGCGGAATA 4372
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4373 CCTTCAGAGTGTGACTGCAAAATGATGCACTGTGCTTACAGAGTGGAGATGGC 4432
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4433 TAGCCCAAGGATGCCAAACTCAGTGGCCCATCTCTCCCTGGCTGTTCTCCAGATGGTACA 4492
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4493 CTGTATATTCAGATCTAGGGAATATCGGATCGGGCTGTGTCAAGAATAAGCCITTA 4552
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4553 CTTAACTCTATGAATCTTATGAAGTTCGGTCTCAACTGATCAAGAACTCTACATCTTT 4612
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4613 GACATCAATGGTACTCCCAATATACCTGTAAGTTTAGTCACTGGTATTACTTTACAT 4672
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4673 TTTAGCTACACCAATGACATGATATTAATCTGCTGTGACAGACAGCAATGGCAACACCTT 4732
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4733 AGAATTAGACGGGACCCCAATCGCATGCCAGTTCGAGTGTCTCTCGTAACCAAGTG 4792
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4793 ATATGGTTGACAAATAGGAACCAATGGATGTTTGAAGGCATGACCTGCTCAAGGACTGAA 4852
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 4853 TTAGTTTGTGTACTTACCATGGCAATAGTGGCTTTTAGCCACTAAAAGTGATGAAT 4912
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 4913 GGATGGCAACAGCTTTTTCGACTATGACAGTGAAGTGTGCTGACAAATGTTACGTTTCCA 4972
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 4973 ACTGGAGTGTTCACAAACCTGATGGGACATGGACAGGCTATCACATGGACATGGAG 5032
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5033 TCATCTAGCCGAGAGAGATGTCAGCATCACTTCAAAATCTGTCTCTCGATCGATTTCTTC 5092

QY 260 TyrThr 261
 Db 5093 TACACC 5098
 RESULT 6
 ABQ82344
 ID ABQ82344 standard; cDNA; 8645 BP.
 XX
 AC ABQ82344;
 DT 17-DEC-2002 (first entry)
 XX
 DE Human NOV15b encoding cDNA SEQ ID NO:37.
 XX
 KW Human; NOVX; cytotatic; neuroprotective; anticonvulsant; cardiovascular;
 KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
 KW antihemetic; antibacterial; immunosuppressive; antiallergic; virucide;
 KW antianaemic; lymphoma; melanoma; neurological disorder; gene therapy;
 KW cancer; leukaemia; Pick's disease; vesicular transport disease; epilepsy;
 KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
 KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
 KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
 KW ulcerative colitis; gastric reaction; duodenal disorder; infection;
 KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
 KW rheumatoid arthritis; gene; chromosome 4; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 151..8316
 FT /*tag= a
 FT /product= "NOV15b"
 XX
 XX WO200262999-A2.
 PN
 PD 15-AUG-2002.
 XX
 PF 31-DEC-2001; 2001WO-US049976.
 XX
 PR 29-DEC-2000; 2000US-0258928P.
 PR 02-JAN-2001; 2001US-0259415P.
 PR 04-JAN-2001; 2001US-0259785P.
 PR 20-FEB-2001; 2001US-0269814P.
 PR 09-MAR-2001; 2001US-0279863P.
 PR 29-MAR-2001; 2001US-0279832P.
 PR 13-APR-2001; 2001US-0279833P.
 PR 18-APR-2001; 2001US-0283889P.
 PR 13-APR-2001; 2001US-0284447P.
 PR 29-MAY-2001; 2001US-0286683P.
 PR 16-AUG-2001; 2001US-0312915P.
 PR 17-AUG-2001; 2001US-0313232P.
 PR 17-SEP-2001; 2001US-0322699P.
 PR 26-NOV-2001; 2001US-0333350P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
 PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
 PI Kudara R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
 PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
 PI Gunther E, Smithson G, Millet I, Macdougall JR;
 XX
 DR WPI: 2002-732706/79.
 DR P-PSDB; ABP53587.
 XX
 XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
 XX associated disorders, such as cancers, neurological disorders, disorders
 XX of vesicular transport, gastrointestinal disorders, and autoimmune
 XX diseases.
 XX
 PS Claim 8; Page 114-117; 444pp; English.

XX
 CC The present invention describes novel human proteins designated NOVX,
 CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
 CC cytotatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
 CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
 CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
 CC protozoacide and antihelminthic activities, and can be used in gene
 CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease selected from NOVX-associated disorder, such as cancers
 CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
 CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
 CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
 CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
 CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
 CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
 CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
 CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
 CC and protozoal infections. The NOVX proteins can be used as immunogens to
 CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
 CC used in chromosome mapping, identifying individuals from minute
 CC biological samples (tissue typing), and in forensic identification of a
 CC located on chromosome 4
 CC
 XX
 SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 2,4e-86 Length: 8645
 Score: 235.50 Matches: 172
 Percent Similarity: 79.01% Conservative: 35
 Best Local Similarity: 67.65% Mismatches: 54
 Query Match: 65.59% Indels: 1
 DB: 6 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x ABQ82344 (1-8645)

QY 1 LyslleAsnArglleArgGlnValThrThrSerGlyGluSerLeuValAlaGlyAla 20
 Db 4489 AAAATTAACCGGATAAGGCAGGTCAACAGATGAGAAATCTCTTAGTGCCTGGAATA 4548
 QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnAspCysAspCysPheSerGlyAspGly 40
 Db 4549 CCTTCAGAGTGTGACTGCCAAAATGATGCCAACTGTGACTGTACACAGATGCAGATGC 4608
 QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
 Db 4609 TACGCCAAGGATGCCAACTAGTCCCACTCTCCCTCCCTGGCTGCTCCAGATGGTACA 4668
 QY 61 LeuTyrValAlaAspLeuGlyAsnIleArglleArgPheIleArgLysAsnLysProphe 80
 Db 4669 CTGTATATTGCAGATCTAGGGAATATCCGGATCCGGGTGTGTCAAGAATAAGCTTTA 4728
 QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
 Db 4729 CTTAACCTATGAATCTTCTATGAAGTTGCGTCTCAACTGATCAAGAACTCTACATCTTT 4788
 QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
 Db 4789 GACATCAATGGTACTCACCATACTAGTAACTTTAGTCACTGGTGATTACTCTTTACAAAT 4848
 QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
 Db 4849 TTTAGCTACAGCAATGACAATGATATTACTGTCTGTGACAGACAGCAATGGCAACACCTTT 4908
 QY 141 AsnValArgArgAspSerThrGlyMetProLeuThrLeuValValProAspGlyGlnVal 160
 Db 4909 AGAATTAGCGGGACCAATTCGCATGCCAGTTTCAGTGTGTCTCTGTATACCAAGT 4968
 QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
 Db 4969 ATATGGTTGACATAGGACAAATGGATGTTTGAAGGCATGACTGCTCAAGGACTGAA 5028

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuAlaThrIysSerAsnGluAsn 200
DB 5029 TTAGTTTCTTTTACTTACCATGCAATAGTGGCTTTTATAGCCACTAAAGTGAAGAACT 5088
QY 201 GlyTTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
DB 5089 GGATGACACACGTTTTTTCATCTATGACATGACAGTGAAGTCTGTGACAAATGTTCAGTTTCCA 5148
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
DB 5149 ACTGGAGTGTCACAACTGCTGGGACATGGCAAGGCTATCATCAGTGGACATTTGAG 5208
QY 241 ThrSerSerIys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
DB 5209 TCATCTAGCGGAGAGAGATGTCAGCATCCTCAATCTGCTCGATCTCTTC 5268
QY 260 TyrThr 261
DB 5269 TACACC 5274
RESULT 7
ABQ82343
ID ABQ82343 standard; cDNA; 8675 BP.
XX AC ABQ82343;
XX XT 17-DEC-2002 (first entry)
XX DE Human NOV15a encoding cDNA SEQ ID NO:35.
XX KW Human; NOVX; cytosolic; neuroprotective; anticonvulsant; cardiovascular;
KW KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW KW antirheumatic; antiarthritic; immunosuppressive; anti-allergic; vitucide;
KW KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW KW rheumatoid arthritis; gene; chromosome 4; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 151..8328
XX FT /*tag= a
XX FT /product= "NOV15a"
XX FT /transl_except= (pos:1249..1251,aa:Ser)
XX PN WO2000262999-A2.
XX PD 15-AUG-2002.
XX PF 31-DEC-2001; 2001WO-US049976.
XX PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX PR (CURA-) CURAGEN CORP.

XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shinkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX WPI; 2002-732706/79.
DR P-PSDB; ABP53586.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX Claim 8; Page 110-112; 444pp; English.
XX The present invention describes novel human proteins designated NOVX,
XX where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
XX cytosolic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
XX antidiabetic, antiinflammatory, antirheumatic, antiarthritic, vitucide,
XX immunosuppressive, antibacterial, antianaemic, antibacterial, fungicide,
XX protozoacide and antihelminthic activities, and can be used in gene
XX therapy. The NOVX proteins, nucleotides or antibodies can be used in the
XX manufacture of a medicament for treating a syndrome associated with a
XX human disease selected from NOVX-associated disorder, such as cancers
XX (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
XX ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
XX ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
XX disease), disorders of vesicular transport (e.g. cystic fibrosis,
XX diabetes mellitus, Grave's disease, or goitre), gastrointestinal
XX disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
XX autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
XX anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
XX and protozoal infections. The NOVX proteins can be used as immunogens to
XX produce antibodies and as vaccines. The NOVX nucleotide sequences may be
XX used in chromosome mapping, identifying individuals from minute
XX biological samples (tissue typing), and in forensic identification of a
XX biological sample. The present sequence encodes human NOV15a, which is
XX located on chromosome 4
XX SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;
Alignment Scores:
Pred. No.: 2,41e-86 Length: 8675
Score: 935.50 Matches: 172
Percent Similarity: 79.01% Conservative: 35
Best Local Similarity: 65.65% Mismatches: 54
Query Match: 67.59% Indels: 1
DB: 6 Gaps: 1
US-10-029-020-14_COPY_1490_1750 (1-261) x ABQ82343 (1-8675)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
DB 4501 AAAATTAAACCGGATAGGACAGTCAACACAGATGAGAAATCTCTTAGTGCGCGAATA 4560
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
DB 4561 CCTTCAGAGTGTGACTGCAAAAATGATGCCAACTGTGACTGTGTACACAGATGGAGATGC 4620
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
DB 4621 TACGCCAAGGATGCCAAACTCAGTCCCATCTCTCCCTGGCTGCTTCCAGATGGTACA 4680
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgGlyAsnLysProPhe 80
DB 4681 CTGTATATGTCAGATCTAGGAAATATCCGATCCGGCTGTGTCTGAAGAAATAAGCCTTTA 4740
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
DB 4741 CTTAACTCTATGAACCTTCTATGAAGTGGTCTCTCCAACTGATCAAGAACTTACATCTTT 4800

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
 Db 4801 GACATCAATGGTACTACCAATATACCTGTAAGTTTACTGCTGATTAACCTTACAAAT 4860
 QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
 Db 4861 TTAGCTACAGCAATGCAATGATATATTACTGCTGTCAGACAGCAATGGCAACCCCTT 4920
 QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
 Db 4921 AGAATTAGACGGGACCCAAATCGCATGCCAGTTCCGAGTGGTGTCTCCTGATAACCAAGTG 4980
 QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
 Db 4981 ATATGGTTGACAAATAGGAACAAATGGATGTTTGAAGGCATGACTGCTCAAGACTGGAA 5040
 QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
 Db 5041 TTAGTTTGTGTTTACTTACCATGGCAATAGTGGCTTTTAGCCACTAAAGTGATGAAACT 5100
 QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
 Db 5101 GGAATGACACACGTTTTTGGACTAAGACAGTGAAGGTGCTGTGACAAATGTTACGTTTCCA 5160
 QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
 Db 5161 ACTGGAGTGGTCACAAACCTGCATGGGACATGGACAGGCTATCACATGGACATTTAG 5220
 QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
 Db 5221 TCATCTAGCCGAGAAGAAGATGTGACATCACTTCAAACTGTCTCGATCGATTTCTTC 5280
 QY 260 TyrThr 261
 Db 5281 TACACC 5286

RESULT 8
 ABK92230
 ID ABK92230 standard; DNA; 12879 BP.
 AC ABK92230;
 XX
 DT 15-AUG-2002 (first entry)
 XX
 DE Prostate cancer-associated DNA sequence #116.
 XX
 KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
 KW gene therapy; gene; ds.
 XX
 OS Mammalia.
 XX
 FN WO200230268-A2.
 XX
 PD 18-APR-2002.
 XX
 EF 12-OCT-2001; 2001WO-US032045.
 XX
 PR 13-OCT-2000; 2000US-00687576.
 PR 08-DEC-2000; 2000US-00733298.
 PR 08-DEC-2000; 2000US-00733742.
 PR 24-JAN-2001; 2001US-0263957P.
 PR 16-MAR-2001; 2001US-0276791P.
 PR 16-MAR-2001; 2001US-0276889P.
 PR 06-APR-2001; 2001US-0281922P.
 PR 24-APR-2001; 2001US-0286214P.
 PR 30-APR-2001; 2001US-00847046.
 PR 04-MAY-2001; 2001US-0288589P.
 XX
 FA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 FI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
 XX
 DR WPI; 2002-471335/50.

P-PSDB; ABG61913.

Detecting a prostate cancer-associated transcript in a cell in a patient,
 useful for diagnosing prostate cancer (PC) or screening modulators of PC,
 by determining if prostate cancer-associated genes are expressed in a
 prostate tissue.

Claim 22; Page 394-397; 436pp; English.

The present invention relates to methods of detecting a prostate cancer-
 associated transcript in a cell from a patient. The method comprises
 contacting a biological sample from the patient with prostate cancer-
 associated polynucleotides (designated PC genes) that selectively
 hybridise to a sequence that is at least 80% identical to them. The
 prostate cancer-associated polynucleotide sequences are differentially
 expressed in prostate tumour tissue or in prostate cancer and are derived
 from the tissues of various organisms such as humans or other mammals
 (e.g. mice, sheep and dogs). The methods of the invention are useful for
 diagnosing and treating prostate cancer in mammals. The prostate cancer-
 associated genes are useful for diagnosing or treating prostate cancer,
 as well as for identifying modulators of prostate cancer or agents that
 inhibit prostate cancer. The nucleic acid sequences are particularly
 useful in gene therapy, as a vaccine or in antisense applications.
 CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
 sequences

Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-------|
| Pred. No.: | 3,16e-82 | Length: | 12879 |
| Score: | 898.00 | Matches: | 162 |
| Percent Similarity: | 78.46% | Conservative: | 42 |
| Best Local Similarity: | 62.31% | Mismatches: | 56 |
| Query Match: | 64.88% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-029-020-14_COPY_1490_1750 (1-261) x ABK92230 (1-12879)

| | | | |
|----|------|--|------|
| QY | 1 | LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla | 20 |
| Db | 4406 | AAAGTAAACCGCATTCAGCAAGTAACCAACCATGGGAGATTCATCATCTGCTGGTGC | 4465 |
| QY | 21 | ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly | 40 |
| Db | 4466 | CCCCTGACTGTGACTGCAAAATTCGACAACTGTGACTGTTTTCAGTGTGCTGTC | 4525 |
| QY | 41 | TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGly | 60 |
| Db | 4526 | TATGCCAAAGATGCAAGATGAAGATGAAGCCCTTCCTCTTAGCAGTGCCTGATGAACC | 4585 |
| QY | 61 | LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe | 80 |
| Db | 4586 | CTCTATGTGGCAGACCTCGGAAATGTTTCAATTCGATCCATCAGCAGAAACCAACCCAC | 4645 |
| QY | 81 | LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe | 100 |
| Db | 4646 | CTGAATGACATCAACATTTATGAGATGCTTTCACCCGCTGATCAGGAATCTACAGTTC | 4705 |
| QY | 101 | AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn | 120 |
| Db | 4706 | ACTGTAATGGAACCCACCTACACCTGAACTGTATACAGGGACTATGTTTATAAC | 4765 |
| QY | 121 | PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal | 140 |
| Db | 4766 | TTACCTTACAATTCGAAAGGTGACTTGGCGCGGATTCACGACGACATGGCAATTCAGT | 4825 |
| QY | 141 | AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal | 160 |
| Db | 4826 | CACATTCGCGTGTATGCGAGCGGAATCCGCTATGGCTTGTGGTCCCTGCGGCAAGTA | 4885 |
| QY | 161 | TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu | 180 |
| Db | 4886 | TACTGGCTGACTATACGCAATGGAGTCTCTGAAAAGAGTGTACGCCCAAGGCTATAAT | 4945 |

XX New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.

PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.

XX The invention relates to isolated polynucleotides which are
XX differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,41e-79 Length: 9695
Score: 863.50 Matches: 162
Percent Similarity: 77.10% Conservative: 40
Best Local Similarity: 61.83% Mismatches: 59
Query Match: 62.39% Indels: 1
Gaps: 7

US-10-029-020-14_COPY_1490_1750 (1-261) x ACC72052 (1-9695)

QY 1 LysileAsnArgileArgGlnValThrSerGlyGluileSerLeuValAlaGlyAla 20
Db 4500 AAGATTAAACGCTTACGCCAGGTAAACCAACCGGGGAGATCTGCCCTTTAGCTGGGCA 4559
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnAspCysAspGlyAspGly 40
Db 4560 GCCTCGACATCGCAGCTGCAAAACGATGTCAATGGCACTGTAATCAGAGATGATGCC 4619
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4620 TAGCGGACTGATGCCATCTTGAATTCGCCATCATCCCTAGCTAGCTCCAGATGGTACC 4679
QY 61 LeuTyrValAlaAspLeuGlyAsnileArgileArgPheileArgLysAsnLysProPhe 80
Db 4680 ATTTACATTGACAGACCTTGGAAATATTCGGATCAGGGCGGTTCAGCAAGAACACCCCTGT 4739
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProLysAspGlnGluLeuTyrLeuPhe 100
Db 4740 CTTAATGCTTCAACACAGTATGAGGCTGCATCCCGGAGAGAGAGTATATGTTTC 4799
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4800 AACGCTGATGATCCACCAATACACTGTGAGCCTGGTGACAGGGGAGTACTTGTACAA 4859
QY 121 PheThrThrGlyAspGlyAspLeuThrLeuLeuThrAspAsnAsnGlyAsnMetVal 140
Db 4860 TTCACATATGATGACATGATGCTCACTGAATTTGATGATGATGATGATGATGATGATG 4919
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
Db 4920 AAGATCCCTCGGAGCAGCAGGATGATGATGATGATGATGATGATGATGATGATGATG 4979
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrGlnGlyHisGlu 180
Db 4980 ATCACCCTCACCGTGGGCACCAATGAGGAGGCTCAAAAGTGTCTCCACACAGAACCTGG 5039
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5040 CTTGTCTCATGACCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5099

QY 201 GlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5100 GGATGGACGACCTTTCTATGACTATGACACCAAGAGCCGCTGACCAAGTGTACGGGCC 5159
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 5160 ACGGGGGTGGTAACCACTGTCACCGGGAATGGAGAAATCTATTACCATTCACATTGAG 5219
QY 241 ThrSerSerLys---AspAspValThrThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5220 AACTCCACCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5279
QY 260 TyrThr 261
Db 5280 TACACA 5285
RESULT 12
AAS14089
ID AAS14089 standard; DNA; 9729 BP.
XX
AC AAS14089;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human FCTR3f DNA sequence.
XX
KW Human; FCTRX; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
KW astrocytoma; congenital neonatal autoimmune thrombocytopaenia; infection;
KW neurological disorder; neurodegenerative disorders; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greenow type I; Corneal dystrophy-lattice type I;
KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW antidiabetic; antiasthmatic; antiinfectivity; antiinflammatory;
KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3a; neurestin-like protein; FCTR3f.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX 5'UTR 1..209 /*tag= b
XX CDS 210..8384 /*tag= a
XX 3'UTR /*product= "Human FCTR3f"
XX 8385..9729 /*tag= c
XX
XX WO200166747-A2.
XX
XX 13-SEP-2001.
XX
XX 05-MAR-2001; 2001WO-US007160.
XX
XX 03-MAR-2000; 2000US-0186592P.
XX PR 03-MAR-2000; 2000US-0186718P.
XX PR 06-MAR-2000; 2000US-0187293P.
XX PR 06-MAR-2000; 2000US-0187294P.
XX PR 17-MAR-2000; 2000US-0190400P.
XX PR 07-APR-2000; 2000US-0196018P.
XX PR 03-JAN-2001; 2001US-0259548P.
XX
XX (CURA-) CURAGEN CORP.
XX Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
XX Macdougall J, Mishra V, Mezes PS, Rastelli L;

XX WPI; 2001-596837/67.
 DR P-PSDB; AAU08681.
 XX
 PT Novel polypeptides designated as FCTR3f polypeptides, useful in detection,
 PT prevention and treatment of a broad range of pathological states.
 XX
 PS Claim 9; Page 37-39; 215pp; English.
 XX
 CC The invention relates to human FCTR3f polypeptides, FCTR1-FCTR7, and the
 CC nucleic acids encoding them. These sequences are useful for the treatment
 CC or prevention of numerous disorders including myelogenous leukaemia,
 CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
 CC autoimmune thrombocytopenia, neurological disorders, neurodegenerative
 CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
 CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
 CC myelodysplastic syndrome, mental health conditions, immunological
 CC disorders, allergy and infection, bronchial asthma, Avellino type
 CC eosinophilia, lung diseases, reproductive disorders, infertility, male
 CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
 CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
 CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
 CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
 CC Corneal dystrophy-Greenou type I, Corneal dystrophy-lattice type I and
 CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
 CC the FCTR3a homologue FCTR3f
 XX
 SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 8.45e-79 Length: 9729
 Score: 863.50 Matches: 162
 Percent Similarity: 77.10% Conservative: 40
 Best Local Similarity: 61.83% Mismatches: 59
 Query Match: 62.39% Indels: 1
 DB: 5 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x AAS14089 (1-9729)

QY 1 LysileAsnArgileAsgGluValThrThrSerGlyGluileSerLeuValalaGlyala 20
 Db 4548 AAGATTAAACCGCTACGCCAGGTACACCAACGAGGAGATCTGCTTTAGCTGGGCA 4607

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
 Db 4608 GCCTCGGACTCGGCTGCAAAACGATGTCATTCGCACTGCTATTCAGGAGATGCC 4667

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
 Db 4668 TACCGGACTGATGCATCTTGAATTCGCCATCATCTTAGCTAGCTCCAGATGGTACC 4727

QY 61 LeuTyrValAlaAspLeuGlyAsnileAargPheileAargLysAsnLysProphe 80
 Db 4728 ATTTCATTCGACACCTTGGAAATATTCGGATACGGCGGCTCAGCAAGAACAGCCTGTT 4787

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProileAspGlnGluLeuTyrLeuphe 100
 Db 4788 CTTAATGCCCTTCAACAGTATGAGCTGCATCCCGGAGAGAGAGGATGATATGTTTC 4847

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
 Db 4848 AAGCCTGATGGCATCCCAACCAATACATCTGAGCCTGCTGAGCGGAGTACTTGTACAAT 4907

QY 121 PheThrTyrThrGlyAspGlyAspPheileThrleuileThrAspAsnAsnGlyAsnMetVal 140
 Db 4908 TTCACATATAGTACTGACATGATGTCACCTGATGATGATGATGATGATGATGATGATG 4967

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTripleuValProAspGlyGlnVal 160
 Db 4968 AAGATCGTCGGGACAGCAGTGGCATGCCCGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 5027

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
 Db 5027

Db 5028 ATCACCCTCACCGTGGGACCAATGAGGCGCTCAAAAGTCGTGTCACACAGAACCTGGAG 5087
 QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
 Db 5088 CTTGGTCTCATGACCTATGATGGCAACACTGGGCTCTGGCCACCAAGCATGAACA 5147
 QY 201 GlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
 Db 5148 GGATGACGACCTTCTCTATGACTATGACACGAGGCGCGCTGACCAACGTCGCGCCCC 5207
 QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
 Db 5208 ACGGGGTGGTAACCACTTCGACCGGAAATGGAGAAATCTATTACCATTCACATGAG 5267
 QY 241 ThrSerSerLys---AspAspValThrThrThrAsnLeuSerAlaSerGlyAlaPhe 259
 Db 5268 AACTCCACCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5327
 QY 260 TyrThr 261
 Db 5328 TACACA 5333

RESULT 13
 ADB32028
 ID ADB32028 standard; cDNA; 9729 BP.
 XX
 AC ADB32028;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human FCTR3f cDNA.
 XX
 KW Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
 KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
 KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
 KW renal cell carcinoma; melanoma; clear cell carcinoma;
 KW granular cell carcinoma; neurological disorder;
 KW neurodegenerative disorder; nerve trauma;
 KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
 KW Gardner syndrome; mental health condition; immunological disorder;
 KW allergy; asthma; lung disease; reproductive disorder; deafness;
 KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
 KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
 KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
 KW plasmodium falciparum infection; Greenou's corneal dystrophy;
 KW lattice corneal dystrophy.
 XX
 OS Homo sapiens.
 XX
 FN US2003087816-A1.
 XX
 PD 08-MAY-2003.
 XX
 PF 05-MAR-2001; 2001US-00800198.
 XX
 PR 03-MAR-2000; 2000US-0186592P.
 XX
 XX (VERM/) VERMET C.
 PA (FERN/) FERNANDES E.
 PA (SHIM/) SHIMKETS R.
 PA (HERR/) HERRMANN J.
 PA (MAJU/) MAJUMDER K.
 PA (MACD/) MACDOUGALL J.
 PA (MISH/) MISHRA V.
 PA (MEZE/) MEZES P S.
 PA (RAST/) RASTELLI L.
 XX
 PI Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
 PI MacDougall J, Mishra V, Mezes PS, Rastelli L;
 XX
 DR WPI; 2003-625633/59.
 DR P-PSDB; ADB32029.
 XX

PT New FCTR3b polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTR3b-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.

XX Claim 9; Page 34-37; 155pp; English.

XX The invention relates to FCTR3b polypeptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTR3b polypeptide, such as colorectal cancer,
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, neurological disorders,
CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
CC conditions, immunological disorders, allergy and infection, asthma, lung
CC diseases, male and female reproductive disorders, deafness, glycoprotein
CC deficiency, desmold tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTR3b polypeptide of the invention.

XX SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

| Alignment Scores: | | Length: 9729 | |
|---|-----------|---|------|
| Pred. No.: | 8, 45e-79 | Matches: | 162 |
| Score: | 863.50 | Conservative: | 40 |
| Percent Similarity: | 77.10% | Mismatches: | 59 |
| Best Local Similarity: | 61.83% | Indels: | 1 |
| Query Match: | 62.39% | Gaps: | 1 |
| DB: | 9 | | |
| US-10-029-020-14_COPY_1490_1750 (1-261) x ADB32028 (1-9729) | | | |
| QY | 1 | LysileAsnArgileArgGlnValThrSerGlyGluileSerLeuValAlaGlyala | 20 |
| DB | 4548 | AAGATTAAACCGTCTACGCCAGGTAAACCAACCGGGAGATCTCCCTTTTATGCTGGGCA | 4607 |
| QY | 21 | ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly | 40 |
| DB | 4608 | GCCTCGGAGCTCGGACTCGGACTCGGAAACAGTGTCAATTCGACACTGCTATTCAGGAGATGATGCC | 4667 |
| QY | 41 | TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu | 60 |
| DB | 4668 | TACGGGACTGATGCCATCTTGAATTCCTCCATCATCTTAGCTGTAGCTCCAGATGGTACC | 4727 |
| QY | 61 | LeuTyrValAlaAspLeuGlyAsnIleArgileArgPheIleArgLysAsnLysProPhe | 80 |
| DB | 4728 | ATTACATTTGACAGACTTGGAAATATTCCGATCAGGGCGGTGACGCAAGAACAGCTGTT | 4787 |
| QY | 81 | LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrIleuPhe | 100 |
| DB | 4788 | CTTAAAGCTTCAACAGTATGAGCTGATCCCGCGGAGACGAGTATATATGTTTC | 4847 |
| QY | 101 | AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrIleuTyrAsn | 120 |
| DB | 4848 | AACGCTGATGGCATCCACCAATACACTGTGAGCTGTGTGACACAGGGGAGTACTTGTACAA | 4907 |
| QY | 121 | PheThrThrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal | 140 |
| DB | 4908 | TTACATATAGTACATGACAAATGATGTCACTGAATTAATGACAAATATGGGAATTCCTCG | 4967 |
| QY | 141 | AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal | 160 |
| DB | 4968 | AAGATCCGTCGGGACAGCAGTGGCATGCCCGTCCCTGCTCATGCTGACCAACAGATC | 5027 |
| QY | 161 | TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu | 180 |
| DB | 5028 | ATCACCCCTCCCGTGGGCAACCAATGAGGCTTCAAAGTCTGTCCACACAGAACCTGGAG | 5087 |

| | | | |
|-----------|---|--|------|
| QY | 181 | LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrIleSerAsnGluAsn | 200 |
| DB | 5088 | CTTGGTCTCATGACCTATGATGGCAACACTGGGCTCTCTGGCCACCAAGAGCGATGAACA | 5147 |
| QY | 201 | GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro | 220 |
| DB | 5148 | GGATGGAGGACTTTCTATGACTATGACCAAGAGCGGCTGACCACTGACGCGCCCC | 5207 |
| QY | 221 | ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu | 240 |
| DB | 5208 | ACGGGGTGGTAAACAGTCTGACCGGAAATGGAGAAATCTATTACCATTCATGACATTGAG | 5267 |
| QY | 241 | ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe | 259 |
| DB | 5268 | AACTCCACCGTATGATGACGCTCACTGTCTATCACCACCACTCTCTTCAGTAGAGGCTCC | 5327 |
| QY | 260 | TyrThr 261 | |
| DB | 5328 | TACACA 5333 | |
| RESULT 14 | | | |
| AAS14085 | | | |
| ID | AAS14085 standard; DNA; 9826 BP. | | |
| AC | AAS14085; | | |
| DT | 18-DEC-2001 (first entry) | | |
| XX | Human FCTR3b DNA sequence. | | |
| DE | Human; FCTR3b; myelogenous leukaemia; carcinoma; melanoma; glioma; ds; | | |
| KW | astrocytoma; congenital neonatal autoimmune thrombocytopenia; infection; | | |
| KW | neurological disorder; neurodegenerative disorders; nerve trauma; | | |
| KW | familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy; | | |
| KW | demyelinating Gardner syndrome; familial myelodysplastic syndrome; | | |
| KW | mental health condition; immunological disorder; allergy; infertility; | | |
| KW | bronchial asthma; Avellino type eosinophilia; lung disease; deafness; | | |
| KW | reproductive disorder; reproductive disorder; glycoprotein Ia deficiency; | | |
| KW | desmold disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide; | | |
| KW | gastric disorders; pancreatic disease; Schistosoma mansoni infection; | | |
| KW | Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes; | | |
| KW | Corneal dystrophy-Groenouw type I; Corneal dystrophy-lattice type I; | | |
| KW | Reis-Bucklers corneal dystrophy; cystostatic; immunosuppressive; | | |
| KW | antiallergic; antiasthmatic; antifertility; antineoplastic; | | |
| KW | antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological; | | |
| KW | gynaecological; antifertility; immunostimulant; auditory; haemostatic; | | |
| XX | Gene therapy; FCTR3b; neurestin-like protein. | | |
| OS | Homo sapiens. | | |
| XX | Key | | |
| FT | 5'UTR | Location/Qualifiers | |
| FT | | 1..279 | |
| FT | | /*tag= b | |
| FT | CDS | 280..8481 | |
| FT | | /*tag= a | |
| FT | | /product= "Human FCTR3b" | |
| FT | | 8482..9826 | |
| FT | 3'UTR | /*tag= c | |
| XX | | WO200166747-A2. | |
| XX | | 13-SEP-2001. | |
| XX | | 05-MAR-2001; 2001WO-US007160. | |
| XX | | 03-MAR-2000; 2000US-0186592P. | |
| PR | | 03-MAR-2000; 2000US-0186719P. | |
| PR | | 06-MAR-2000; 2000US-0187293P. | |
| PR | | 06-MAR-2000; 2000US-0187294P. | |
| PR | | 07-MAR-2000; 2000US-0190400P. | |
| PR | | 17-APR-2000; 2000US-0196018P. | |
| PR | | 03-JAN-2001; 2001US-0259548P. | |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 302.635 Seconds

(without alignments)
2440.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_1250_1400

Perfect score: 782

Sequence: 1 YIRIFPSONVTNILELRNK.....LLGSNDLTARPLSCDSYMD 151

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519
-DB=Published_Applications_NA -QFM=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020@cgn1.1.2156@runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|----|-------------|
|------------|-------|--------------------|----|-------------|

ALIGNMENTS

RESULT 1

US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

| 1 | 782 | 100.0 | 8354 | 13 | US-10-383-201-43 | Sequence 43, Appl |
|----|-------|-------|-------|----|----------------------|--------------------|
| 2 | 782 | 100.0 | 8354 | 13 | US-10-029-020-13 | Sequence 13, Appl |
| 3 | 759.5 | 97.1 | 8355 | 13 | US-10-383-201-55 | Sequence 55, Appl |
| 4 | 759.5 | 97.1 | 8438 | 13 | US-10-042-865-1 | Sequence 1, Appl |
| 5 | 625 | 79.9 | 8409 | 9 | US-09-808-602-79 | Sequence 79, Appl |
| 6 | 625 | 79.9 | 8409 | 10 | US-09-800-198-67 | Sequence 67, Appl |
| 7 | 619 | 79.2 | 8473 | 17 | US-10-038-854-39 | Sequence 39, Appl |
| 8 | 619 | 79.2 | 8487 | 17 | US-10-038-854-41 | Sequence 41, Appl |
| 9 | 619 | 79.2 | 8645 | 17 | US-10-038-854-37 | Sequence 37, Appl |
| 10 | 619 | 79.2 | 8675 | 17 | US-10-038-854-35 | Sequence 35, Appl |
| 11 | 613 | 78.4 | 6560 | 9 | US-09-808-602-76 | Sequence 76, Appl |
| 12 | 613 | 78.4 | 6560 | 10 | US-09-800-198-64 | Sequence 64, Appl |
| 13 | 613 | 78.4 | 9058 | 16 | US-10-144-194A-79 | Sequence 79, Appl |
| 14 | 613 | 78.4 | 9695 | 16 | US-10-144-194A-81 | Sequence 81, Appl |
| 15 | 613 | 78.4 | 9729 | 9 | US-09-808-602-12 | Sequence 12, Appl |
| 16 | 613 | 78.4 | 9729 | 10 | US-09-800-198-12 | Sequence 12, Appl |
| 17 | 613 | 78.4 | 9826 | 9 | US-09-808-602-7 | Sequence 7, Appl |
| 18 | 613 | 78.4 | 9826 | 10 | US-09-800-198-7 | Sequence 7, Appl |
| 19 | 610 | 78.0 | 8689 | 9 | US-09-808-602-78 | Sequence 78, Appl |
| 20 | 610 | 78.0 | 8689 | 10 | US-09-800-198-66 | Sequence 66, Appl |
| 21 | 610 | 78.0 | 8797 | 9 | US-09-808-602-74 | Sequence 74, Appl |
| 22 | 610 | 78.0 | 8797 | 10 | US-09-800-198-62 | Sequence 62, Appl |
| 23 | 610 | 78.0 | 8797 | 10 | US-09-800-198-65 | Sequence 65, Appl |
| 24 | 610 | 78.0 | 8797 | 10 | US-10-072-012-143 | Sequence 143, Appl |
| 25 | 602.5 | 77.0 | 8575 | 13 | US-10-072-012-143 | Sequence 143, Appl |
| 26 | 534.5 | 68.4 | 12880 | 16 | US-10-295-027-927 | Sequence 927, Appl |
| 27 | 413 | 52.8 | 247 | 15 | US-10-029-386-16574 | Sequence 16574, A |
| 28 | 413 | 52.8 | 557 | 15 | US-10-029-386-2874 | Sequence 2874, Ap |
| 29 | 211 | 27.0 | 751 | 13 | US-10-027-632-160821 | Sequence 160821, |
| 30 | 211 | 27.0 | 751 | 16 | US-10-027-632-160821 | Sequence 160821, |
| 31 | 111 | 14.2 | 789 | 13 | US-10-424-599-25144 | Sequence 25144, A |
| 32 | 101 | 12.9 | 1482 | 9 | US-09-938-842A-1949 | Sequence 1949, Ap |
| 33 | 101 | 12.9 | 1482 | 11 | US-09-938-842A-1949 | Sequence 1949, Ap |
| 34 | 98 | 12.5 | 1480 | 13 | US-10-424-599-30579 | Sequence 30579, A |
| 35 | 98 | 12.5 | 2276 | 13 | US-10-424-599-30579 | Sequence 30579, A |
| 36 | 93 | 11.9 | 1719 | 13 | US-10-425-114-20338 | Sequence 20338, A |
| 37 | 89 | 11.4 | 1851 | 17 | US-10-437-963-88447 | Sequence 88447, A |
| 38 | 88.5 | 11.3 | 1431 | 13 | US-10-424-599-132367 | Sequence 132367, |
| 39 | 88.5 | 11.3 | 1925 | 17 | US-10-437-963-11176 | Sequence 11176, A |
| 40 | 85 | 10.9 | 2550 | 16 | US-10-104-047-72 | Sequence 72, Appl |
| 41 | 83 | 10.6 | 1735 | 13 | US-10-425-114-707 | Sequence 707, App |
| 42 | 82.5 | 10.5 | 2061 | 13 | US-10-424-599-13785 | Sequence 13785, A |
| 43 | 82 | 10.4 | 4377 | 9 | US-09-815-242-9570 | Sequence 9570, Ap |
| 44 | 81.5 | 10.4 | 450 | 15 | US-10-029-386-16526 | Sequence 16526, A |
| 45 | 81.5 | 10.4 | 555 | 15 | US-10-029-386-2826 | Sequence 2826, Ap |

;; PRIOR FILING DATE: 2002-09-23
;; PRIOR APPLICATION NUMBER: 10/051,874
;; PRIOR FILING DATE: 2002-01-16
;; PRIOR APPLICATION NUMBER: 60/366,928
;; PRIOR FILING DATE: 2002-03-22
;; PRIOR APPLICATION NUMBER: 10/055,877
;; PRIOR FILING DATE: 2002-01-22
;; NUMBER OF SEQ ID NOS: 155
;; SOFTWARE: CuraseqList version 0.1
;; SEQ ID NO 43
;; LENGTH: 8354
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 9,15e-108 Length: 8354
Score: 782.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-383-201-43 (1-8354)

QY 1 TyrlleAArgArgilePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3782 TACATTAGAGGATCTCCCTCTGGAAATGTCCACCAATCTCTAGAGCTGAGGAATAAA 3841
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3842 GATTTCCAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGT 3901
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3902 GGGGCGGCTCTCTCTTCAGAGCAACAGCCGGGGGCTCTTTAAATCAAGTCCACCTGTG 3961
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3962 GTGGTGAAGACCTTGTCAAGAACTCTGAGTGGTTCGGGGACAGGTGACCATGCTC 4021
QY 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4022 CCCTTTGATGACACTCGCTCGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACC 4081
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4082 CCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGACCATGATC 4141
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4142 AGACGCATCGATCAGATGGATCATCTCCACCTGCTCGGCTCTAATGATCTCACATCA 4201
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4202 GCCCGGCCACTCAGCTGTGATTCTGTCTATGGAT 4234

RESULT 2
US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590

;; PRIOR FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: 60/257,314
;; PRIOR FILING DATE: 2000-12-20
;; PRIOR APPLICATION NUMBER: 60/311,613
;; PRIOR FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: 60/315,617
;; PRIOR FILING DATE: 2001-08-29
;; PRIOR APPLICATION NUMBER: 60/307,506
;; PRIOR FILING DATE: 2001-07-24
;; PRIOR APPLICATION NUMBER: 60/322,358
;; PRIOR FILING DATE: 2001-09-14
;; PRIOR APPLICATION NUMBER: 60/294,075
;; PRIOR FILING DATE: 2001-05-29
;; PRIOR APPLICATION NUMBER: 60/288,153
;; PRIOR FILING DATE: 2001-05-02
;; NUMBER OF SEQ ID NOS: 190
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 13
;; LENGTH: 8354
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-029-020-13

Alignment Scores:
Pred. No.: 9,15e-108 Length: 8354
Score: 782.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-029-020-13 (1-8354)

QY 1 TyrlleAArgArgilePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3782 TACATTAGAGGATCTCCCTCTGGAAATGTCCACCAATCTCTAGAGCTGAGGAATAAA 3841
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3842 GATTTCCAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGT 3901
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3902 GGGGCGGCTCTCTCTTCAGAGCAACAGCCGGGGGCTCTTTAAATCAAGTCCACCTGTG 3961
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3962 GTGGTGAAGACCTTGTCAAGAACTCTGAGTGGTTCGGGGACAGGTGACCATGCTC 4021
QY 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4022 CCCTTTGATGACACTCGCTCGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACC 4081
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4082 CCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGACCATGATC 4141
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4142 AGACGCATCGATCAGATGGATCATCTCCACCTGCTCGGCTCTAATGATCTCACATCA 4201
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4202 GCCCGGCCACTCAGCTGTGATTCTGTCTATGGAT 4234

RESULT 3
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A

; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 2,6e-104 | Length: | 8355 |
| Score: | 759.50 | Matches: | 151 |
| Percent Similarity: | 96.79% | Conservative: | 0 |
| Best Local Similarity: | 96.79% | Mismatches: | 0 |
| Query Match: | 97.12% | Indels: | 5 |
| DB: | 13 | Gaps: | 2 |

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-383-201-55 (1-8355)

| | | | |
|----|------|--|------|
| Qy | 1 | TyrileArgArgilePheProSerGlyAsnValThrAsnIleLeuGluLeu-----Arg | 18 |
| Db | 3754 | TACATTAGAGGATCTTCCCTCTGGAATGTCACCAATCTTAGAGCTGAGGTCAGA | 3813 |
| Qy | 19 | AsnLysAspPheArgHisSerHisSerProAlaHisLysTyrTyrIleuAlaThrAspPro | 38 |
| Db | 3814 | AATAAGATTTCAGACATAGTCACAGTCAGCACACAATACTACTCTGGCCACAGACCCC | 3873 |
| Qy | 39 | MetSerGlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSer | 58 |
| Db | 3874 | ATGAGTGGGGCGCTTCTTCTTGACAGCAACAGCGCGGGTCTTTTANAATCAAGTCC | 3933 |
| Qy | 59 | ThrValValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGln | 78 |
| Db | 3934 | ACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAGTGGTGGTGGGACAGGTGACCAG | 3993 |
| Qy | 79 | CysLeuProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeu | 98 |
| Db | 3994 | TGCTCCCTCCCTTTGTATGACACTCGCTCGCGGATGCTGGGAAGGCCACAGAAGCCACACTC | 4053 |
| Qy | 99 | ThrAsnProArg-----GlyIleThrValAspLysPheGlyLeuIleTyrPheVal | 115 |
| Db | 4054 | ACCAATCCAGGGGTCCCGCCAGGCAATTACAGTGGACAAAGTTTGGGTGATCTACTTCGTG | 4113 |
| Qy | 116 | AspGlyThrMetIleArgArgileAspGlnAsnGlyIleIleSerThrLeuLeuGlySer | 135 |
| Db | 4114 | GATGGCACCATGATCAGACGATCATGAGATGGATCATCTCCACCTCTCGGCTCT | 4173 |
| Qy | 136 | AsnAspLeuThrSerAlaArgProLeuSerCysAspSerValMetAsp | 151 |

Db 4174 AATGATCTCATCATCAGCCCGGCCACTCAGCTGTGATTCTCTCATGGAT 4221

RESULT 4

US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 2,64e-104 | Length: | 8438 |
| Score: | 759.50 | Matches: | 151 |
| Percent Similarity: | 96.79% | Conservative: | 0 |
| Best Local Similarity: | 96.79% | Mismatches: | 0 |
| Query Match: | 97.12% | Indels: | 5 |
| DB: | 13 | Gaps: | 2 |

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-042-865-1 (1-8438)

| | | | |
|----|------|---|------|
| Qy | 1 | TyrileArgArgilePheProSerGlyAsnValThrAsnIleLeuGluLeu-----Arg | 18 |
| Db | 3764 | TACATTAGAGGATCTTCCCTCTGGAATGTCACCAATCTTAGAGCTGAGGTCAGA | 3823 |

QY 19 AsnLysAspPheArgHisSerProAlaHisLysTyrThrLeuAlaThrAspPro 38
 Db 3824 AATAAAGATTTCAGACATAGTCACAGTCCAGCACACAAATATCTACTGCCACAGACCC 3883
 QY 39 MetSerGlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysLysSer 58
 Db 3884 ATGAGTGGGGCGCTCTTCCTTCTGACACACAGCCGGGGCTCTTTAAATCAAGTCC 3943
 QY 59 ThrValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGln 78
 Db 3944 ACTGTGCTGGTGAAGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGTGAC 4003
 QY 79 CysLeuProPheAspThrArgCysGlyAspGlyValLysAlaThrGluAlaThrLeu 98
 Db 4004 TGCCTCCCTTTGATGACACTCGCTCGGGGATGGTGGGAAGCCACAGAACCCACATC 4063
 QY 99 ThrAsnProArg-----GlyIleThrValAspLysPheGlyLeuIleTyrPheVal 115
 Db 4064 ACCAATCCAGGGGTCCCCAGGCGATTCAGTGGACAAAGTTTGGGTGATCTACTTCGTG 4123
 QY 116 AspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySer 135
 Db 4124 GATGGCACCATGATCAGACGATCGATCAGATGGGATCATCTCCACCTGCTCGGCTCT 4183
 QY 136 AsnAspLeuThrSerAlaArgProLeuSerCysAspSerValMetAsp 151
 Db 4184 AATGATCTCATCATCAGCCCGGCGCACTCAGCTGTGATTCGTGTATGGAT 4231

RESULT 5
 US-09-808-602-79
 ; Sequence 79, Application US/09808602
 ; Patent No. US20020155115A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vernet, Corine A
 ; APPLICANT: Fernandes, Elma
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Herrmann, John L
 ; APPLICANT: Majumder, Kumud
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Mezes, Peter S
 ; APPLICANT: MacDougall, John
 ; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 15966-697 CIP
 ; CURRENT FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 09/800,198
 ; PRIOR FILING DATE: 2001-03-05
 ; PRIOR APPLICATION NUMBER: 60/186,596
 ; PRIOR FILING DATE: 2000-03-03
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 79
 ; LENGTH: 8409
 ; TYPE: DNA
 ; ORGANISM: Gallus gallus
 US-09-808-602-79
 Alignment Scores:
 Pred. No.: 1.14e-83 Length: 8409
 Score: 625.00 Matches: 113
 Percent Similarity: 90.07% Conservative: 23
 Best Local Similarity: 74.83% Mismatches: 15
 Query Match: 79.92% Indels: 0
 DB: Gaps: 9
 US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-808-602-79 (1-8409)
 QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
 Db 3856 TACATTCGGCGTATCTTCCATCCAGGATGTGACTAGCATATTGGAGCTGAGAAATAA 3915
 QY 21 AspPheArgHisSerProAlaHisLysTyrThrLeuAlaThrAspProMetSer 40
 Db 3915 GAGTTTAAACATAGCAACAATCTCTGCTCACAAATCTATCTGCGCGTGGACCCCGTTTCG 3975
 QY 41 GlyAlaValPheLeuSerAspSerArgValPheLysLysSerThrVal 60

Db 3916 GAGTTTAAACATAGCAACAATCTCTGCTCAAAATACTATCTGGCGCTGGACCCCGTTTCG 3975
 QY 41 GlyAlaValPheLeuSerAspSerArgValPheLysLysSerThrVal 60
 Db 3976 GGCCTCCCTGCTAGTATCAGACACACAGCCGCGGATATACAAAGTCAAAATCTCTTACT 4035
 QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
 Db 4036 GGCACGAAAGACCTGGCTGTAATCTGAAGTGGTAGCGGGACTGGAGCAATCCCTG 4095
 QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyValAlaThrGluAlaThrLeuThrAsn 100
 Db 4096 CCCTTTGATGAAGCCAGATGTGGAGATGGAGGAAAGAGTGGACGCACACCTTAATGAGT 4155
 QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
 Db 4156 CCTCGAGGAATTCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4215
 QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 Db 4216 CGAAAAGTGGATCAGAAATGGAATATATCAACTCTCTGGGCTCCAATGACCTTAACCTGCC 4275
 QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
 Db 4276 GTCGACCTCTAAGCTGTGATTCAGCATGGAT 4308

RESULT 6
 US-09-800-198-67
 ; Sequence 67, Application US/09800198
 ; Publication No. US20030087816A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vernet, Cornie AM
 ; APPLICANT: Fernandes, Elma
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Herrmann, John L
 ; APPLICANT: Majumder, Kumud
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Mezes, Peter S
 ; APPLICANT: Rastelli, Luca
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 15966-697
 ; CURRENT APPLICATION NUMBER: US/09/800,198
 ; CURRENT FILING DATE: 2001-03-05
 ; PRIOR APPLICATION NUMBER: 60/186,596
 ; PRIOR FILING DATE: 2000-03-03
 ; NUMBER OF SEQ ID NOS: 98
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 67
 ; LENGTH: 8409
 ; TYPE: DNA
 ; ORGANISM: Gallus gallus
 US-09-800-198-67
 Alignment Scores:
 Pred. No.: 1.14e-83 Length: 8409
 Score: 625.00 Matches: 113
 Percent Similarity: 90.07% Conservative: 23
 Best Local Similarity: 74.83% Mismatches: 15
 Query Match: 79.92% Indels: 0
 DB: Gaps: 10
 US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-800-198-67 (1-8409)
 QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
 Db 3856 TACATTCGGCGTATCTTCCATCCAGGATGTGACTAGCATATTGGAGCTGAGAAATAA 3915
 QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrThrLeuAlaThrAspProMetSer 40
 Db 3916 GAGTTTAAACATAGCAACAATCTCTGCTCACAAATCTATCTGCGCGTGGACCCCGTTTCG 3975
 QY 41 GlyAlaValPheLeuSerAspSerArgValPheLysLysSerThrVal 60

```
Db 3976 GGCTCCCTGTACGTATCAGACACCAACGCGGATATACAAAGTCAAACTCTTACT 4035
QY 61 ValVallyAspLeuVallyAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 4036 GGCACGAAGACCTGGCTGGTAATCTGAAGTGGTAGCGGGGACTGGAGCAATGCCTG 4095
QY 81 ProPheAspThrArgCysGlyAspGlyGlyAlaThrGluAlaThrLeuThrAsn 100
Db 4096 CCCTTGTAGACGACAGATGTGGAGATGGAGGAAGCAGTGGACCAACCTTAATGAGT 4155
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4156 CCTCGAGGAATGCGAGTGGATAAGTATGACATCATTTTGTGATGCCACTATGATT 4215
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4216 CGAAAGTGGATCAGAAATGGAATTATATCAACTCTGCTGGGTCCAATGACCTAATGCTGC 4275
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4276 GTCCGACCTCTAAGCTGTGATTCAGCATGGAT 4308
```

RESULT 7

```
US-10-038-854-39
; Sequence 39, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,899

; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 8473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-39

Alignment Scores:
Pred. No.: 9,65e-83 Length: 8473
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-038-854-39 (1-8473)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3600 TACGTGCGCGGATATTCCTTCTGAAATGTAAACAGTGTCTTAGAACTAAGAAATAAA 3659
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3660 GATTTAGACATAGCAGCAACCCAGCTCATAGTACTACCTTGCACCGATCCATCAGC 3719
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3720 GGAGATCTGTAGTTCTGTGACACAAACACCCGACAGATTATTCGCCCAAGTCACCTAGC 3779
QY 61 ValVallyAspLeuVallyAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3780 GGGGCAAAAGACTTGACTTAAATAATGCAGAGTCGTCGACGGGACAGGGGAGCAATGCCTT 3839
QY 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 3840 CCGTTTGACGAGCGGAGATGTGGGATGGAGGAGGCCGTGGAGGCCACACTCATGAGT 3899
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 3900 CCCAAAGGAATGGCAGTTGATAAGAAATGATTAATCTACTTTTGTGATGGAACCATGATT 3959
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 3960 AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTGACTTCA 4019
QY 141 AlaArgProLeuSerCysAspSerValMet 150
Db 4020 GCCAGACCTTTAACTTGTGCACACCATG 4049

RESULT 8
US-10-038-854-41
; Sequence 41, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
```

```

; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-038-854-41

Alignment Scores:
Pred. No.: 9,686-83      Length: 8487
Score: 619.00           Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16%      Indels: 0
DB: 17                  Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-038-854-41 (1-8487)

Qy 1 TyrIleArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnIly 20
Db 3596 TACGTGGCGGAGATTCCTCTTGGAAATGTATACAAAGTGTCTTAGAACTAAGAAATAAA 3655

Qy 21 AspPheArgHisSerHisSerProAlaHisIlySerTyrIleuAlaThrAspProMetSer 40
Db 3656 GAITTTAGACATACACACCCAGCTCATAGATACCTTTCGCAACGGATCCAGTCCAG 3715

Qy 41 GlyAlaValPheLeuSerAspSerAsnSerArgValPheIlyIleLysSerThrVal 60
Db 3716 GGAGATCTGTAGCTTCTCTGACACAAACCCGCGAGATTTATCGCCCAAGTCACTTACG 3775

Qy 61 ValIlyLysAspLeuValLysAsnSerGluValValIleAlaGlyThrGlyAspGlnCysLeu 80
Db 3776 GGGGCAAAAGACTTGTACTTAAATGCAAGATCTCGAGGAGGAGGAGGAGCAATGCTT 3835

Qy 81 ProPheAspAspThrArgCysGlyAspGlyGlyLeuAlaThrGluAlaThrLeuThrAsn 100
Db 3836 CCGTTTGACGAGCGAGATGTGGGATGGAGGAGCGCGTGGAGCCACACATCATGAGT 3895

Qy 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 3896 CCAAAGGAATGGCAGTGTGATAAGATGGATTAATCTACTTTGTGATGGAACCATGATT 3955

Qy 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 3956 AGGAAAGTTGACCAAAATGAATCATATCAACTCTTCTGGGCTCTTACGATTTGACTTCA 4015

Qy 141 AlaArgProLeuSerCysAspSerValMet 150
Db 4016 GCCAGACCTTTAACTTGTGACACCCAGCATG 4045

RESULT 9
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
```

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 411

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 37

; LENGTH: 8645

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-038-854-37

Alignment Scores:

Pred. No.: 9,97e-83 Length: 8645
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-038-854-37 (1-8645)

Qy 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3772 TACGTGGCGGATATTCCTTCTGGAATGTAAACAGTGTCTTAAAGTAAATAAA 3831
Qy 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3832 GATTTAGACATAGCAGCAACCCAGCTCATAGATCTACTTGCACCGATCCAGTCACG 3891
Qy 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3892 GGAGATCTGTAGTCTTCTGACACAAACCCGACGAATTTATCGCCCAAGTCACCTTACG 3951
Qy 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3952 GGGGCAAAAGACTTGTACTTAAATAATCGCAAGTCGTCGACGGGACAGGGGACAAATGCCTT 4011
Qy 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4012 CCGTTTGACGGCGAGATGTGGGATGGAGGAGCGGTGGAAGCCACACTCATGAGT 4071
Qy 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4072 CCCAAAGGATGGCAGTTGTATAAGAATGCAATTAATCTACTTTGTTGATGAACCATGAT 4131
Qy 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4132 AGGAAGTTGACCAAAATGGAATCATATCAATCTTCTGGGCTCTAACGATTTGACTTCA 4191
Qy 141 AlaArgProLeuSerCysAspSerValMet 150
Db 4192 GCCAGACCTTTAACTTGTGACACCAGCATG 4221

RESULT 10
US-10-038-854-35
; Sequence 35, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spyttek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shinkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderina, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G

; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35

Alignment Scores:
Pred. No.: 1e-82 Length: 8675
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-038-854-35 (1-8675)

Qy 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3784 TACGTGGCGGATATTCCTTCTGGAATGTAAACAGTGTCTTAAAGTAAATAAA 3843
Qy 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3844 GATTTAGACATAGCAGCAACCCAGCTCATAGATCTACTTGCACCGATCCAGTCACG 3903
Qy 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3904 GGAGATCTGTAGTCTTCTGACACAAACCCGAGATTTATCGCCCAAGTCACCTTACG 3963
Qy 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3964 GGGGCAAAAGACTTGTACTTAAATAATGCAGAGTCGTGCGAGGACAGGGGACGATGCTT 4023
Qy 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4024 CCGTTTGACGGCGAGATGTGGGATGGAGGAGCGCGTGAAGCCACACTCATGAGT 4083


```
Db 1022 CGBAGGTTGACCAGATGGAAATCATCTCCACCCCTGCTGGGCTCCATGACTCACTGCC 1081
Qy 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 1082 GTCCGGCCGCTGAGCTGTGATTCAGCATGGAT 1114
RESULT 13
US-10-144-194A-79
; Sequence 79, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 79
; LENGTH: 9058
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (44)..(7699)
US-10-144-194A-79
Alignment Scores:
Pred. No.: 8,95e-82 Length: 9058
Score: 613.00 Matches: 112
Percent Similarity: 89.40% Conservative: 23
Best Local Similarity: 74.17% Mismatches: 16
Query Match: 78.39% Indels: 0
DB: 16 Gaps: 0
US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-144-194A-79 (1-9058)
Qy 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnIlys 20
Db 3146 TACATCCGACGCATCTTCCCTCTCGAAATGTGACCAAGCATCTTGGAGTTACGAAATAAA 3205
Qy 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3206 GAGTTTAAACATAGCAACACCCAGCACACAGAGGAAATCTACCGCGTCAAGTCTCTGAGT 3265
Qy 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3266 GGCTCGCTCTACGTGTCCGACACCAACAGCAGGAGAAATCTACCGCGTCAAGTCTCTGAGT 3325
Qy 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3326 GGAACCAAGACCTGGCTGGGAAATTCGGAAGTTGTGGCAGGAGCGGGAGAGACAGTGCTTA 3385
Qy 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 3386 CCCTTTGATGAGCCCGCTGGGGATGGAGGAGCCATAGATGCACCTCATGAGC 3445
Qy 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 3446 CCGAGAGGTATTGCGATAGACAAGAAATGGGCTCATGTACTTTGTCGATGCCACCATGATC 3505
Qy 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 3506 CGGAAGGTGACCAAGATGGAAATCATCTCCACCCCTGCTGGGCTCCAATGACCTCACTGCC 3565
Qy 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 3566 GTCCGGCCGCTGAGCTGTGATTCAGCATGGAT 3598
RESULT 14
US-10-144-194A-81
; Sequence 81, Application US/10144194A
; Publication No. US20030215809A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 9695
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (435)..(8336)
US-10-144-194A-81
Alignment Scores:
Pred. No.: 9,98e-82 Length: 9695
Score: 613.00 Matches: 112
Percent Similarity: 89.40% Conservative: 23
Best Local Similarity: 74.17% Mismatches: 16
Query Match: 78.39% Indels: 0
DB: 16 Gaps: 0
US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-144-194A-81 (1-9695)
Qy 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnIlys 20
Db 3783 TACATCCGACGCATCTTCCCTCTCGAAATGTGACCAAGCATCTTGGAGTTACGAAATAAA 3842
Qy 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3843 GAGTTTAAACATAGCAACACCCAGCACACCAAGTACTACTTGGCAGTGGACCCCGTGTC 3902
Qy 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3903 GGCTCGCTCTACGTGTCCGACCAACAGCAGGAGAAATCTACCGCGTCAAGTCTCTGAGT 3962
Qy 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3963 GGAACCAAGACCTGGCTGGGAAATTCGGAAGTTGTGGCAGGAGCGGGAGAGACAGTGCTTA 4022
Qy 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4023 CCCTTTGATGAGCCCGCTGGGGATGGAGGAGGCCATAGATGCACACCTCATGAGC 4082
Qy 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4083 CCGAGAGGTATTGCGATAGACAAGAAATGGGCTCATGTACTTTGTCGATGCCACCATGATC 4142
Qy 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4143 CGGAAGGTGACCAAGATGGAAATCATCTCCACCCCTGCTGGGCTCCAATGACCTCACTGCC 4202
Qy 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4203 GTCCGGCCGCTGAGCTGTGATTCAGCATGGAT 4235
RESULT 15
US-09-808-602-12
; Sequence 12, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herriman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
```

```
FILE REFERENCE: 15966-697 CIP
CURRENT APPLICATION NUMBER: US/09/808,602
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/800,198
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
PRIOR FILING DATE: 2000-03-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 9729
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (210)..(8381)
US-09-808-602-12

Alignment Scores:
Pred. No.: 1e-81 Length: 9729
Score: 613.00 Matches: 112
Percent Similarity: 89.40% Conservative: 23
Best Local Similarity: 74.17% Mismatches: 16
Query Match: 78.39% Indels: 0
DB: Gaps: 9

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-808-602-12 (1-9729)

QY 1 TyrlleArgArgilePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3831 TACATCCGACGCACTCTCCCTCTCGAAATGTGACCAGCACTCTGGAGTTACGAAATAAA 3890
QY 21 AspPheArgHisSerHisProAlaHisLysIleTyrLeuAlaThrAspProMetSer 40
Db 3891 GAGTTTAACATAGCACACACCAGCACACAAGTACTACTTGGCAGTGGACCCCGTCTCC 3950
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3951 GGCTCGCTCTACGTGCCGACACCAACAGCAGGAGAATCTACCGCGTCAAGTCTCTGAGT 4010
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 4011 GGAACCAAGACCTGGCTGGGAATTCGAAAGTTGTGGCAGGACGGGAGACGAGTGCTA 4070
QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4071 CCCTTTGATCAAGCCCGCTCGCGGGATGGAGGAGAGGCCATAGATGCAACCCCTGATGAGC 4130
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4131 CCGAGAGGTATTGCAGTAGACAAGAATGGGCTCATGTACTTGTTCGATGCCACCATGATC 4190
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4191 CGAAGGTTACCAAGATGGAATCATCTCCACCTCTGGGCTCCAAATGACCTCAGTCGCC 4250
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4251 GTCCGCGCGCTGAGCTGTGATTCAGCATGGAT 4283
```

Search completed: August 14, 2004, 19:18:15
Job time : 333.635 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 44.6731 Seconds

(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_1250_1400

Perfect score: 782

Sequence: 1 YIRIFPSGVNTNILELNK.....LLGSNDLSARPLSCDSVMD 151

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=Issued Patents NA -QEMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020.cgn2_1.1.258 @runat_06082004_112217_29301 -NCPUS=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|----|--------------------|
| c 1 | 90.5 | 11.6 | 4403765 | 3 | US-09-103-840A-2 |
| c 2 | 90.5 | 11.6 | 4411529 | 3 | US-09-103-840A-1 |
| c 3 | 83 | 10.6 | 4403765 | 3 | US-09-103-840A-2 |
| c 4 | 83 | 10.6 | 4411529 | 3 | US-09-103-840A-1 |
| c 5 | 81 | 10.4 | 9607 | 4 | US-08-961-527-80 |
| c 6 | 80 | 10.2 | 978 | 4 | US-09-134-000C-790 |
| c 7 | 79.5 | 10.2 | 2736 | 4 | US-09-235-451-3 |
| c 8 | 78.5 | 10.0 | 14042 | 3 | US-08-652-877-85 |
| c 9 | 78.5 | 10.0 | 14044 | 3 | US-08-652-877-89 |
| c 10 | 78.5 | 10.0 | 14080 | 3 | US-08-652-877-87 |
| c 11 | 78.5 | 10.0 | 14083 | 3 | US-08-476-515A-83 |
| c 12 | 78.5 | 10.0 | 14086 | 3 | US-08-652-877-83 |

| | | | | | | |
|------|------|-----|------|---|----------------------|--------------------|
| c 13 | 76.5 | 9.8 | 6943 | 4 | US-09-453-702B-213 | Sequence 213, Appl |
| c 14 | 76 | 9.7 | 2073 | 4 | US-09-134-001C-2101 | Sequence 2101, Ap |
| c 15 | 75 | 9.6 | 2797 | 4 | US-09-453-702B-244 | Sequence 244, Appl |
| c 16 | 74.5 | 9.5 | 1302 | 4 | US-09-252-991A-14317 | Sequence 14317, A |
| c 17 | 74.5 | 9.5 | 1479 | 4 | US-09-252-991A-14286 | Sequence 14286, A |
| c 18 | 73 | 9.3 | 2028 | 2 | US-08-933-750C-73 | Sequence 73, Appl |
| c 19 | 73 | 9.3 | 2028 | 3 | US-09-234-613-73 | Sequence 73, Appl |
| c 20 | 73 | 9.3 | 4351 | 4 | US-09-060-299-24 | Sequence 24, Appl |
| c 21 | 73 | 9.3 | 4351 | 4 | US-09-402-923A-24 | Sequence 24, Appl |
| c 22 | 73 | 9.3 | 4843 | 4 | US-09-060-299-2 | Sequence 2, Appli |
| c 23 | 73 | 9.3 | 4843 | 4 | US-09-402-923A-2 | Sequence 2, Appli |
| c 24 | 73 | 9.3 | 4915 | 4 | US-09-060-299-28 | Sequence 28, Appl |
| c 25 | 73 | 9.3 | 4915 | 4 | US-09-402-923A-28 | Sequence 28, Appl |
| c 26 | 73 | 9.3 | 4959 | 4 | US-09-060-299-38 | Sequence 38, Appl |
| c 27 | 73 | 9.3 | 5022 | 4 | US-09-402-923A-38 | Sequence 38, Appl |
| c 28 | 73 | 9.3 | 5022 | 4 | US-09-060-299-32 | Sequence 32, Appl |
| c 29 | 73 | 9.3 | 5098 | 4 | US-09-402-923A-32 | Sequence 32, Appl |
| c 30 | 73 | 9.3 | 5098 | 4 | US-09-060-299-1 | Sequence 1, Appli |
| c 31 | 73 | 9.3 | 5098 | 4 | US-09-402-923A-1 | Sequence 1, Appli |
| c 32 | 73 | 9.3 | 5125 | 4 | US-09-060-299-26 | Sequence 26, Appl |
| c 33 | 73 | 9.3 | 5125 | 4 | US-09-402-923A-26 | Sequence 26, Appl |
| c 34 | 73 | 9.3 | 5162 | 4 | US-09-060-299-33 | Sequence 33, Appl |
| c 35 | 73 | 9.3 | 5162 | 4 | US-09-402-923A-33 | Sequence 33, Appl |
| c 36 | 73 | 9.3 | 5166 | 4 | US-09-060-299-23 | Sequence 23, Appl |
| c 37 | 73 | 9.3 | 5166 | 4 | US-09-402-923A-23 | Sequence 23, Appl |
| c 38 | 73 | 9.3 | 5263 | 4 | US-09-060-299-31 | Sequence 31, Appl |
| c 39 | 73 | 9.3 | 5263 | 4 | US-09-402-923A-31 | Sequence 31, Appl |
| c 40 | 72.5 | 9.3 | 885 | 4 | US-09-446-301A-3 | Sequence 3, Appli |
| c 41 | 72.5 | 9.3 | 885 | 4 | US-09-099-932-3 | Sequence 3, Appli |
| c 42 | 72.5 | 9.3 | 1607 | 4 | US-09-446-301A-16 | Sequence 16, Appl |
| c 43 | 72.5 | 9.3 | 1607 | 4 | US-09-099-932-50 | Sequence 50, Appl |
| c 44 | 72 | 9.2 | 1599 | 4 | US-09-060-299-37 | Sequence 37, Appl |
| c 45 | 72 | 9.2 | 1599 | 4 | US-09-402-923A-37 | Sequence 37, Appl |

ALIGNMENTS

RESULT 1
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:

Pred. No.: 1.65e+03
Score: 90.50
Percent Similarity: 45.00%
Best Local Similarity: 27.00%
Query Match: 11.57%
DB: 3
Matches: 27
Conservative: 18
Mismatches: 22
Indels: 33
Gaps: 4

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-103-840A-2 (1-4403765)

Mon Aug 16 09:01:06 2004

```

; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 1.94e+04 4411529
Score: 83.00 28
Percent Similarity: 39.81% Conservative: 15
Best Local Similarity: 25.93% Mismatches: 29
Indels: 36
Query Match: 10.61% Gaps: 5
DB: 3

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-103-840A-1 (1-4411529)

QY 30 HisLysTyrLeuAlaThrAspProMetSerGlyAlaValPheLeuSerAspSerAsn 49
Db 3961444 CACCCCTATGATGTCGGGTGACGGT---GCTGGCACCCTGCTACGTACCGACGCGC 3961500
QY 50 SerArgArgValPheLysLysLysSerThrValValValLysAspLeuValLysAsnSer 59
Db 3961501 CACAACTCGCGTGGTGGCGTTG----- 3961521
QY 70 GluValValAlaGlyThrGlyAspGlnCys---LeuProPheAspAspThrArgCysGly 88
Db 3961522 -----ACCGGGGGTTCGGCCACCGCGGTGCACCTCCATTCGCCGAT----- 3961563
QY 89 AspGlyGlyAlaThrGluAlaThrLeuThrAsnProArgGlyLeuThrValAspLys 108
Db 3961564 -----CTCAGCTTCCCGCGGTGTGCGGTGGACCGC 3961596
QY 109 PheGlyLeuLeuThrPheValAspGlyThrMetIleArgArgIleAspGlnAsnGlyIle 128
Db 3961597 GACGATAGCTCTATGTGGCCGAT-----CTGAACAACATCGCGGTG 3961638
QY 129 IleSerThrLeuLeuGlySerAsn 136
Db 3961639 CTGAAGCTGGCGCGCGCTCGAAT 3961662

RESULT 5
US-08-961-527-80
; Sequence 80, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: Hp Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders

```

```

; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-80

Alignment Scores:
Pred. No.: 3.09 9607
Score: 81.00 39
Percent Similarity: 41.56% Conservative: 25
Best Local Similarity: 25.32% Mismatches: 53
Query Match: 10.36% Indels: 37
DB: 4 Gaps: 7

US-10-029-020-14_COPY_1250_1400 (1-151) x US-08-961-527-80 (1-9607)

QY 19 AsnLysAspPheArgHisSerHis-----SerProAlaHisLysTyrTyrLeu 34
Db 5204 MATCGAATCTCGGTGTCAGTGAACATGCCAACGACGACCATTCCTCAAGGCTCATTTT 5263
QY 35 AlaThr-----AspProMetSerGlyAlaValPheLeuSerAspSerAsnSerArg 52
Db 5264 CGAAGCGCTAATGAATGATGGTGAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 5323
QY 53 ValPheLysLysSerThrValValValLysAspLeuValLysAsnSerGluValVal 72
Db 5324 CTGGTT---ATTGAACAACCAATGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 5380
QY 73 AlaGlyThrGlyAspGlnCysLeuProPheAspAspThrArgCysGlyAspGlyGlyLys 92
Db 5381 -----AAGGGTGACTGTGTATAGCCTTTCATCGACAAG-----GCTGAA 5419
QY 93 AlaThrGluAlaThrLeuThr-----AsnProArgGly 103
Db 5420 GAAACAGTTGCTGAGTTGACCTATAAGAAAGCTTTTGGATTTATGGAATCCGCTGCCA 5479
QY 104 IleThrValAsp-----LysPheGly 110
Db 5480 GATATGTTGATTGCGGATTGAAAGAAATTAACATCCATCTACTGGGGGATGGATTGCT 5539
QY 111 LeuIleTyrPheValAspGlyThrMetIleArgArgIleAspGlnAsnGly-IleIleSe 130
Db 5540 GTGATTTATCTGGCATCGCAGATGCTGTGCAACGTTCTATGAACCGGGTTATTTGGTT 5599

RESULT 6
US-09-134-000C-790
; Sequence 790, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055, 778
; FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 790
; LENGTH: 978
; TYPE: DNA

```

ORGANISM: Enterococcus faecalis
US-09-134-000C-790

Alignment Scores:

Pred. No.: 0.129 Length: 978
Score: 80.00 Matches: 37
Percent Similarity: 39.01% Conservative: 18
Best Local Similarity: 26.24% Mismatches: 58
Query Match: 10.23% Indels: 28
DB: 4 Gaps: 6

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-134-000C-790 (1-978)

QY 9 GlyAsnValThrAsnIle--LeuGluLeuArgAsnLysAspPheArgHisSerHisSer 27
DB 289 GCAATGCGATGAACATACCTCGGCCATTGCAATTTGAAGCATTTGATAATCAAT--- 345
QY 28 ProAlaHisLysTy-TyrLeuAlaThrAspProMetSerGlyAlaValPheLeuSerAsp 47
DB 346 -----ATTATGGGAACAATCCTGTTTCTGCGATGTTGTGTTATTGAC 390
QY 48 -----SerAsnSerArgValPheLysIleLysSerThrValValLys 63
DB 391 GGACGCCAGCTAAATAATGAGTATCGAAATATAATAAATAAACCCTCAAGTCTCTGAT 450
QY 64 AspLeuValLysAsnSerGluValVal-----AlaGlyThrGlyAspGlnCys 79
DB 451 GATTACGCTTCTATGAGAGAAGTATTTATCGCAGATATTCACGAGTTTAAAGAGAGC 510
QY 80 LeuProPheAspThrArgCysGlyAspGlyLys-----AlaThrGlu 95
DB 511 TTGCTTTTCTGATTTAATTTAATCGATGTTGTAAGTCAAGTCGATGTTGCCAA 570
QY 96 AlaThrLeuThrAsnProArgGlyLleThrValAspLysPheGlyLeuIleTy-PheVal 115
DB 571 GATGTGTTAGCAATCAACTAGGAGTGCATATCTCTGAGTGGATTA----- 618
QY 116 AspGlyThrMetIleArgGlyLleAspGlnAsnGlyLleLleSerThrLeuLeuGlySer 135
DB 619 -----GCCAAGACGACAAACGACAAACGAGTGAACATTTTATTGTTGCCA 663
QY 136 Asn 136
DB 664 AAC 666

RESULT 7

US-09-235-451-3/c
; Sequence 3, Application US/09235451
; GENERAL INFORMATION:
; APPLICANT: Julius, David J.
; APPLICANT: Caterina, Michael J.
; APPLICANT: Brake, Anthony J.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING
; TITLE OF INVENTION: CAPSAICIN RECEPTOR AND CAPSAICIN RECEPTOR-RELATED
; FILE REFERENCE: 9076/084CIP
; CURRENT APPLICATION NUMBER: US/09/235,451
; PRIOR FILING DATE: 1999-01-22
; PRIOR FILING DATE: 1998-01-22
; PRIOR FILING DATE: 1998-01-22
; PRIOR FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2736
; TYPE: DNA
; ORGANISM: R. rattus
; FEATURE:
US-09-235-451-3

Alignment Scores:

Pred. No.: 0.735 Length: 2736

Score: 79.50 Matches: 42
Percent Similarity: 41.38% Conservative: 18
Best Local Similarity: 28.97% Mismatches: 79
Query Match: 10.17% Indels: 6
DB: 4 Gaps: 5
US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-235-451-3 (1-2736)
QY 7 ProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLysAspPheArgHisSerHis 26
DB 453 CCCTCTGGAATGCTGACTCCATGCGGGGGGGTTCCTGCTTCCCTTGTTCACCTCAGCAT 394
QY 27 SerProAlaHisLysTy-TyrLeuAlaThrAspProMetSerGlyAlaValPheLeuSer 46
DB 393 TCCCTCTTCTATCTCATCGGAAGTCTCCAGCTCGAAAGCTGGGGGGCTGGAGCTGAAG 334
QY 47 AspSerAsnSerArgValPheLysIleLysSerThrValValLysAspLeuVal 66
DB 333 TCATCTCTGAGAGAGAGGATGCTCTTACAGCTGAACCTCTGTAGAGATGAGATGTTA 274
QY 67 LysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeuProPheAspThr-Ar 86
DB 273 AGGTTCTCTCTCATACCGCAGGGCAGGCTGGCTGGGACCCCGCAGCTGAAGCCAG 214
QY 86 GYSGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsnProArgGlyIleThrVa 106
DB 213 GTGT---AGCGGGGCAAAA---CTCAGGCTAGGATCACCAGCGCTAGCGG---GCTCT 163
QY 106 lasPhysPheGlyLeuIleTy-PheVal---AspGlyThrMetIleArgArgIleAspG 125
DB 162 CCGCAGCGGGGACTGAGAGAGAGAGTGGCGCCGCGGGGTGGCGCTCAACGCTGACAG 103
QY 125 nAsnGlyLleLleSerThrLeuLeuGlySerAsnAspLeuThrSerAlaArgProLeuSe 145
DB 102 TAATGGGGCTGTGGAAACCTCCGCGCGCGCTCCGAGCTTCTCTGCTGCTCCACCCCTC 43
QY 145 r---CysAspSer 148
DB 42 CACCTGTGCTCT 30

RESULT 8

US-08-652-877-85
; Sequence 85, Application US/08652877
; Patent No. 6187548
; GENERAL INFORMATION:
; APPLICANT: Akersstrom, Goran
; APPLICANT: Juhlin, Claes
; APPLICANT: Raak, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjalms, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thereof and DNA Encoding Same
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd., 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.5.1
; SOFTWARE: Word 6.0 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,877
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203


```
QY 45 LeuSerAspSerAsnSerArgValPheLysIleLys-----SerThr 59
Db 2333 TTTTCAGATATGCAAAACACATGATTTTAAACAAAGATTGATGCCACAGAGAA 2392
QY 60 ValValValLysAspLeuValLysAsnSerGluValValala-----73
Db 2393 ATTCTCGCAGCTAACACAGGCTGAAATGTTGAAAGTTTGGCTTTGATTGGATTCAAAG 2452
QY 74 -----GlyThrGlyAspGlnCysLeuProPheAspThrArgCysGlyAspGly 90
Db 2452 AATCTCTATTGGACAGACTCTCATTTACAGAGTATCAGTGTCTAGAGGCTAGCTGATAAA 2512
QY 91 GlyLysAlaThrGluAlaThr-----LeuThrAsnProArgGlyIleThrValAspLysPhe 109
Db 2513 ACGAGACGCACGCTAGTTCAGTATTTAAATAACCCACGCTCGGTGCTATCTCTTTT 2572
QY 110 ----GlyLeuIleThrPheValAsp-GlyThrMetIleArgArgIleAspGlnAsnGlyI 128
Db 2573 GCGGGTATCTATCTTCTACTGATGTTCCGCTCTGCTAAATTTATGAGATGAGT 2632
QY 128 eileSerThrLeuGlySerAsnAspLeuThrSer 140
Db 2633 G-----ACGGATCTCACTCT 2648

RESULT 10
US-08-652-877-87
; Sequence 87, Application US/08652877
; Patent No. 6187548
; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Juhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjalms, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thereof and DNA Encoding Same
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd., 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.5.1
; SOFTWARE: Word 6.0 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,877
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203
; FILING DATE: 22-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,314
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A1355E-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 87:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 14080 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Kidney
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 105..14072
; US-08-652-877-87

Alignment Scores:
Pred. No.: 12.6 Length: 14080
Score: 78.50 Matches: 32
Percent Similarity: 41.35% Conservative: 23
Best Local Similarity: 24.06% Mismatches: 54
Query Match: 10.04% Indels: 24
DB: 3 Gaps: 5

US-10-029-020-14_COPY_1250_1400 (1-151) x US-08-652-877-87 (1-14080)
QY 25 SerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPhe 44
Db 2313 TCGGGATCCTCTCTTTCTTTGCGGATGATTTTTCGCCGCCAGCAGCATCTCTTT 2372
QY 45 LeuSerAspSerAsnSerArgValPheLysIleLys-----SerThr 59
Db 2373 TTTTCAGATATGCAAAACACATGATTTTAAACAAAGATTGATGCCACAGAGAA 2432
QY 60 ValValValLysAspLeuValLysAsnSerGluValValala-----73
Db 2433 ATTCTCGCAGCTAACACAGGCTGAAATGTTGAAAGTTTGGCTTTGATTGGATTCAAAG 2492
QY 74 -----GlyThrGlyAspGlnCysLeuProPheAspThrArgCysGlyAspGly 90
Db 2493 AATCTCTATTGGACAGACTCTCATTTACAGAGTATCAGTGTCTAGAGGCTAGCTGATAAA 2552
QY 91 GlyLysAlaThrGluAlaThr-----LeuThrAsnProArgGlyIleThrValAspLysPhe 109
Db 2553 ACGAGACGCACAGTATGTTCCAGTATTTAAATAACCCACGCTCGGTGCTATCTCTTT 2612
QY 110 ----GlyLeuIleThrPheValAsp-GlyThrMetIleArgArgIleAspGlnAsnGlyI 128
Db 2613 GCGGGTATCTATCTTCTACTGATGTTCCGCTCTGCTAAATTTATGAGATGAGT 2672
QY 128 eileSerThrLeuGlySerAsnAspLeuThrSer 140
Db 2673 G-----ACGGATCTCACTCT 2688

RESULT 11
US-08-476-515A-83
; Sequence 83, Application US/08476515A
; Patent No. 6239270
; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Juhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjalms, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thereof and DNA Encoding Same
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Martin Savitzky
; STREET: Rhone-Poulenc Rorer Inc., 500 Arcola Rd.;
; STREET: 3C43,
```



```

; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Compaq PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 7.0 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,515A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/SE94/00483
; FILING DATE: 24-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9301764-8
; FILING DATE: 24-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A1355D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 83:
; LENGTH: 14083 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..14071
; US-08-476-515A-83
;
Alignment Scores:
Pred. No.: 12.6 Length: 14083
Score: 78.50 Matches: 32
Percent Similarity: 41.35% Conservative: 23
Best Local Similarity: 24.06% Mismatches: 54
Query Match: 10.04% Indels: 24
DB: 3 Gaps: 5
;
US-10-029-020-14_COPY_1250_1400 (1-151) x US-08-476-515A-83 (1-14083)
Qy 25 SerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPhe 44
Db 2315 TCGGGAAATCTCTCTTTCTTCGGGATGATTTTGAGCCGACGACGACATCTTT 2374
Qy 45 LeuSerAspSerAsnArgValPheLysIleLys-----SerThr 59
Db 2375 TTTTCAGATATGTCAAACACATGATTTTAAAGCAAAAGATTGATGGCAGGAAGAA 2434
Qy 60 ValValValLysAspLeuValLysAsnSerGluValValala----- 73
Db 2435 ATTCTCGCAGCTAACAGGGTGGAATAATGTTGAAGTTTGGCTTTTGGATTGCAAAAG 2494
Qy 74 -----GlyThrGlyAspGlnCysLeuProPheAspThrArgCysGlyAspGly 90
Db 2495 AATCTCTATTGGACGACTCTCATACAGAGATATCAGTGTATCAGGCTAGCTGATAA 2554
Qy 91 GlyLysAlaThrGluAlaThr---LeuThrAsnProArgGlyIleThrValAspLysPhe 109
Db 2555 ACGAGACGCACAGTAGTTTCAGTATTAAATAACCCACGCGTGGTGTAGTTTCATCTTTT 2614
;
Qy 110 ---GlyLeuIleTyrPheValAsp-GlyThrMetIleArgIleAspGlnAsnGlyI 128
Db 2615 GCCGGTATCTATTCTTCACGTATGGTTCGCTCTGCTAAATATGAGAGCATGGAGT 2674
Qy 128 eileSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 2675 G-----ACGGATCTCACTCT 2690
;
RESULT 12
US-08-652-877-83
; Sequence 83, Application US/08652877
; Patent No. 6187548
; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Juhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjaln, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thereof and DNA Encoding Same
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd., 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.5.1
; SOFTWARE: Word 6.0 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,877
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203
; FILING DATE: 22-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,314
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A1355E-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14086 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..14074
; US-08-652-877-83
;
Alignment Scores:
Pred. No.: 12.6 Length: 14086
```

Score: 78.50 Matches: 32
Percent Similarity: 41.35% Conservative: 23
Best Local Similarity: 24.06% Mismatches: 54
Query Match: 10.04% Indels: 24
DB: 3 Gaps: 5

US-10-029-020-14_COPY_1250_1400 (1-151) x US-08-652-877-83 (1-14086)

QY 25 SerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPhe 44
Db 2315 TCGGGAAATCTTCTTTCTTGTGGGATGATTTTGACCCGAGGACGACATCTTT 2374

QY 45 LeuSerAspSerAsnSerArgValPheLysLysLys-----SerThr 59
Db 2375 TTTTCAGATATGCAAAACACATGATTTTAAAGCAAAGATTGATGCGCAGGAGAGAA 2434

QY 60 ValValValLysLeuValLysAsnSerGluValValala-----73
Db 2435 ATTCTCGCAGCTAACAGGGGTGGAATAATGTTGAAAGTTTGCTTTTGATTGGATTTCAAAG 2494

QY 74 -----GlyThrGlyAspGlnCysLeuProPheAspAspThrArgCysGlyAspGly 90
Db 2495 AATCTCTATTGGACAGACTCTCATTACAGAGTATCATGCTCATAGGCTAGCTGATAA 2554

QY 91 GlyLysAlaThrGluAlaThr---LeuThrAsnProArgGlyLysLeuThrValAspLysPhe 109
Db 2555 ACGAGCGCACAGTAGTTCAGTATTTAAATAACCCACGGTGGTGTAGTTCATCCTTT 2614

QY 110 ---GlyLeuLeuTyrPheValAsp-GlyThrMetIleArgArgIleAspGlnAsnGlyI 128
Db 2615 GCCGGGTATCTATTCTTCACTGATGTTCCGCTCGCTAAATAATTATGAGCATGGAGT 2674

QY 128 eileSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 2675 G-----ACGGATCTCACTCT 2690

RESULT 13

US-09-453-702B-213/c
; Sequence 213, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000

; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 213:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6943
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 213:
US-09-453-702B-213

Alignment Scores:

Pred. No.: 8,22 Length: 6943
Score: 76.50 Matches: 36
Percent Similarity: 37.34% Conservative: 23
Best Local Similarity: 22.78% Mismatches: 68
Query Match: 9.78% Indels: 31
DB: 4 Gaps: 7

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-453-702B-213 (1-6943)

QY 13 AsnIleLeuGluLeuArgAsnLysAspPheArgHisSerHisSerProAlaHisLysTyr 32
Db 987 AATTATAGCGTTCTGTGTAACCTGTGATTCAAACATAGCAATCGCGCTTATCAGAAATA 928

QY 33 TyrLeuAlaThrAspProMetSerGlyAlaValPheLeuSerAspSerAsnSerArg 52
Db 927 TATTATACCGACACCTGCTCCGGTATGTTTATAGCACCACCGCACTGCTTGCT 868

QY 53 ValPheLysLysSerThrValValValLysAspLeuValLysAsnSerGluValVal 72
Db 867 TTTTACTATCTTAAACGATATGTCGATGTCGGAACAAATAATATCT-----GTG 820

QY 73 AlaGlyThrGlyAspGlnCysLeuProPheAsp-----83
Db 819 CTAATGCGGGGTATGCGGAGTTCTTTTGAACATGTTTCCAACACGAGCACTACAACA 760

QY 84 AspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsnProArgGly 103
Db 759 GATCACACTGT---CAGGGAACAACTACAGCGTTGCGGTGAACTGGA 703

QY 104 IleThrValAsp---LysPheGlyLeuIleTyrPheValAspGlyThrMetIleArg 122
Db 702 GCAGATCGCAAGATTTCAITTCGTATTAAACGTTCAATTAATGGAACGTTAGTAATACCT 643

QY 123 IleAspGlnAsnGlyIleIle-----SerThrLeuLeuGly-----134
Db 642 ATCCCGATATTGCAITGCTGTATGCCAATATCCAGCACCACCGCCCTGTGAGCGG 583

QY 135 -----SerAsnAspLeuThrSerAlaArgProLeuSerCys 146
Db 582 ATTGCAAAAGTTTCGAATTTTCAGGCAGTTTGACCGCA-----CCACAGTCTTGT 535

RESULT 14

US-09-134-001C-2101
; Sequence 2101, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lyvin Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2101
; LENGTH: 2073
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis

```
US-09-134-001C-2101
Alignment Scores:
Pred. No.: 1.52 Length: 2073
Score: 76.00 Matches: 32
Percent Similarity: 33.86% Conservative: 11
Best Local Similarity: 25.20% Mismatches: 32
Query Match: 9.72% Indels: 52
DB: 4 Gaps: 5

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-134-001C-2101 (1-2073)
QY 20 LysAspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPro--- 38
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
196 AAGATTAAATAAACAATGCGAAAGACGAAAGAAATATTTTAGCTAGTGACCCCTGAC 255
QY 39 -----MetSerGlyAlaValPheLeuSerAspSerAsn 49
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
256 CGTGAAGGTGAAGCGATTGCTTGGCAATTATCAAAATTTTAGAATTAGAGATAGCAAA 315
QY 50 SerArgArgValPheLysIleLysSerThrValValValLysAspLeuValLysAsnSer 69
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
316 GAAATAGAGTAGTA-----TTTAATGAAATTTACAAAAGATGCTGTTAAAGATAGT 366
QY 70 GluValValAlaGlyThrGlyAspGlnCysLeuProPheAspAspThrArgCysGlyAsp 89
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
366 ----- 366
QY 90 GlyGlyLysAlaThrGluAlaThrLeuThrAsnProArgGlyIleThrValAspLysPhe 109
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
367 -----TTTAAGCATCCTCGTGTATTGAATGGAT----- 396
QY 110 GlyLeuIleTyrPheValAspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIle 129
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
397 -----TTAGTTGACGCGCAACACAGCAGCTCGT-----ATTTTA 429
QY 130 SerThrLeuLeuGlySerAsn 136
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
430 GATAGACTCGTGGTTATAT 450

RESULT 15
US-09-453-702B-244/c
; Sequence 244, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Buriand, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
```

```
;
;
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 244:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2797
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 244:
US-09-453-702B-244
Alignment Scores:
Pred. No.: 3.34 Length: 2797
Score: 75.00 Matches: 38
Percent Similarity: 36.99% Conservative: 26
Best Local Similarity: 21.97% Mismatches: 59
Query Match: 9.59% Indels: 50
DB: 4 Gaps: 8

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-453-702B-244 (1-2797)
QY 2 IleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnIysAsp 21
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1541 ATTCTAATTTCTCCCGACATAATCTGACC-----CCAAATGAGAGC 1497
QY 22 PheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGly 41
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1496 TGGAAAGATTTCGTGAGACATCGCGCGCTTTTACGTTGCTCAGAGAGTATCACTGGA 1437
QY 42 -----AlaValPheLeuSerAspSerAsnSerArgVal 53
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1436 GATAAGAGTGAACGCTCTTACGCGAAACTCTCTATCTCTGATGCGAGGAATTACAATGACC 1377
QY 54 PheLysIleLysSerThrValValValLysAspLeuValLysAsnSerGluValValAla 73
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1376 TTCAGATA-----GGTGATGTTGTCATC 1353
QY 74 GlyThrGlyAspGlnCysLeuProPheAspAspThr-----ArgCysGlyAspGlyGly 91
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1352 AATACTATTAGCACTGCTATCTCTGAAGATGCAACGGGTCAACGGTGTATCGAAGGGTTG 1293
QY 92 LysAlaThrGluAlaThrLeuThrAsnProArgGlyIleThrValAspLysPheGlyLeu 111
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1292 AATTAGCAGAGATGGATTTAACCGAC-----ATAGACTTGTGAAAATGCGGCTA 1242
QY 112 -----IleTyrPheValAspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIle 129
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1241 AGGAATGTCAATTTT---AATGGCAGCATTTCTTGAATAATGCCAAGTCTTCGCGACCATC 1185
QY 130 -----SerThrLeu 132
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1184 TGTGAAGCGGTGGATTTTACCGATTGTGATCTGCGTAATGAGAAATCGAAATGCCTCA 1125
QY 133 LeuGlySerAsnAspLeuThrSerAlaArgProLeuSer 145
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
1124 TTAGAAAATAGATGATTTTCGTAAAGTTCGCCACTTGACT 1086

Search completed: August 14, 2004, 21:33:24
Job time : 5053.67 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 1770.9 Seconds
(without alignments) 2546.273 Million cell updates

Title: US-10-029-020-14_COPY_1250_1400
Perfect score: 782
Sequence: 1 YIRRIIPSGNVTNILELRNK.....LLGSNDLTSARPLSCDSVMD 151

Scoring table: BLOSUM62

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

| Maximum DB seq length: | 200000000 |
|------------------------|-----------|
| Maximum DB seq length: | 200000000 |

Post-processing: Minimum Match 0%

code processing. Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```

-MODEL=frame_p2n.model -DRV=xlh
-Q/cn21 US1PT0.spool/US10029020/runat_06082004.112216 29287/app_query.fasta.1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOCPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blslog62 -TRANS=human0.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=p2o -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020.@CN1.1.13135 @runat_06082004.112216 29287 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : EST: *

1: em_estba:*
2: em_estbun:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_huv
18: em_gss_luv
19: em_gss_pln
20: em_gss_vrt
21: em_gss_vrt
22: em_gss_mam
23: em_gss_nam
24: em_gss_pro
25: em_gss_rdg
26: em_gss_rpd
27: em_gss_vrl
28: qb_gss1:*

29: gb gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|-------|--------|----|----------|-------------|
| | | | | | | | |
| C | 1 | 727.5 | 93.0 | 6246 | 29 | AY413475 | Homo sapi |
| | 2 | 712.5 | 91.1 | 3190 | 29 | AY413477 | Mus muscu |
| | 3 | 645.5 | 82.5 | 770 | 14 | CB524534 | UI-M-FY0- |
| | 4 | 618 | 79.0 | 623 | 14 | CF538162 | UI-M-GH0- |
| | 5 | 603 | 77.1 | 640 | 14 | CF744767 | UI-M-GV0- |
| | 6 | 588 | 75.2 | 757 | 14 | CD351739 | UI-M-GH0- |
| | 7 | 564.5 | 72.2 | 5069 | 29 | AY405421 | Pan trogl |
| | 8 | 564.5 | 72.2 | 5094 | 29 | AY405420 | Homo sapi |
| | 9 | 563.5 | 72.1 | 5087 | 29 | AY405422 | Mus muscu |
| | 10 | 546.5 | 69.9 | 778 | 13 | BU705631 | UI-M-FO0- |
| C | 11 | 540.5 | 69.1 | 820 | 13 | BU226585 | 603800545 |
| | 12 | 499 | 63.8 | 801 | 14 | CD803099 | UI-M-GV0- |
| | 13 | 484 | 61.9 | 692 | 14 | CF530821 | UI-M-GH0- |
| | 14 | 459 | 58.7 | 724 | 14 | CD352566 | UI-M-GH0- |
| | 15 | 435 | 55.6 | 789 | 14 | CD803315 | UI-M-GV0- |
| | 16 | 408 | 52.2 | 975 | 14 | CD326096 | AGENCOURT |
| | 17 | 401 | 51.3 | 428 | 28 | AZ040891 | AGENCOURT |
| | 18 | 398 | 50.9 | 443 | 10 | BF950887 | RC3-NN118 |
| | 19 | 339 | 43.4 | 351 | 13 | C82439 | RC2439 rabb |
| | 20 | 339 | 43.4 | 351 | 13 | C83295 | C83295 rabb |
| C | 21 | 297 | 38.0 | 784 | 29 | BX144542 | Danio rer |
| | 22 | 285 | 36.4 | 1039 | 12 | BM804083 | AGENCOURT |
| | 23 | 282 | 36.1 | 452 | 13 | BY263427 | BY263427 |
| | 24 | 281 | 35.9 | 319 | 28 | AQ552763 | RCPC1-11-4 |
| | 25 | 264 | 33.8 | 854 | 14 | CF745232 | UI-M-GV0- |
| | 26 | 256 | 32.7 | 843 | 13 | BU262834 | 603818972 |
| | 27 | 242 | 30.9 | 5970 | 29 | AY413476 | Pan trogl |
| | 28 | 223 | 28.5 | 733 | 29 | CNS05671 | Tetraodon |
| | 29 | 221.5 | 28.3 | 1005 | 29 | CNS04860 | Tetraodon |
| | 30 | 211 | 27.0 | 1071 | 29 | CNS05982 | Tetraodon |
| C | 31 | 210 | 26.9 | 640 | 28 | AQ896853 | HS 3144 A |
| | 32 | 202 | 25.8 | 497 | 28 | BZ902852 | CH240 23D |
| | 33 | 202 | 25.8 | 601 | 13 | BX504880 | BZ504880 |
| | 34 | 184 | 23.5 | 662 | 29 | CE057449 | DKFZp686G |
| | 35 | 183 | 23.4 | 723 | 29 | AG030576 | tigr-gss- |
| | 36 | 167 | 21.4 | 645 | 28 | AZ413372 | Pan trogl |
| | 37 | 166 | 21.2 | 1461 | 28 | CZ289028 | CH261-170 |
| | 38 | 163 | 20.8 | 794 | 14 | CB518529 | UI-M-GH0- |
| | 39 | 141 | 18.0 | 360 | 9 | AV191089 | AV191089 |
| | 40 | 139 | 17.8 | 601 | 13 | BX509644 | DKFZp686N |
| C | 41 | 138 | 17.6 | 602 | 9 | AL710647 | DKFZp686M |
| | 42 | 111.5 | 14.3 | 595 | 28 | BH329569 | CH230-106 |
| | 43 | 109 | 13.9 | 532 | 12 | BI418126 | LJ_NEST23a |
| | 44 | 105.5 | 13.5 | 610 | 13 | BU027852 | QHG8D16_Y |
| | 45 | 100 | 12.8 | 361 | 14 | F14403 | ATTS5325 AC |

ALIGNMENTS

| | |
|------------|--|
| RESULT 1 | |
| AY413475 | |
| LOCUS | 6246 bp DNA linear GSS 12-DEC-2003 |
| DEFINITION | Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence. |
| ACCESSION | AY413475 |
| VERSION | AY413475.1 GI:39769437 |
| KEYWORDS | GSS. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 6246) |

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 6246)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..6246
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>6246
/locus_tag="HCM4903"
ORIGIN
Alignment Scores:
Pred. No.: 1.63e-79 Length: 6246
Score: 727.50 Matches: 144
Percent Similarity: 95.36% Conservative: 0
Best Local Similarity: 95.36% Mismatches: 0
Query Match: 93.03% Indels: 7
DB: 29 Gaps: 1
US-10-029-020-14_COPY_1250_1400 (1-151) x AY413475 (1-6246)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 1699 TACATTGAGGAGGATCTCCCTCTGGAAATGTCCACCAATCTAGAGCTG----- 1749
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrIleuAlaThrAspProMetSer 40
Db 1750 -----AGTCACAGTCCAGCACACAAATACTACCTGCCACAGACCCCATGAGT 1797
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 1798 GGCGCGCGTCTCTCTCTGACAGCAACAGCCGCGGGTCTTTAAATCAAGTCCACTGTG 1857
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 1858 GTGTGAGGACCTTGTCAAGACTCTGAGTGTGTTGGCGGACAGGTGACCAAGTGCCTC 1917
QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 1918 CCCTTTGATGACACTCGCTCGCGGGATGTGGGAAGGCCACAGAACCACTACCAAT 1977
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 1978 CCAGGGGCAATTACAGTGGAAGATTTGGGCTGATCTACTCTGTGGATGGCAACATG 2037
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 2038 AGACGATCATCAGAAATGGATCATCTCCACCTGCTCGGCTTAATGATCTACATCA 2097
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 2098 GCCCGGCCACTCAGCTGTGATTTCTGTATGGAT 2130
RESULT 2
AY413477 3190 bp DNA linear GSS 12-DEC-2003
LOCUS
DEFINITION Mus musculus HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,

genomic survey sequence.
ACCESSION AY413477
VERSION AY413477.1 GI:39769439
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3190)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3190)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..3190
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>3190
/locus_tag="HCM4903"
ORIGIN
Alignment Scores:
Pred. No.: 4.65e-78 Length: 3190
Score: 712.50 Matches: 138
Percent Similarity: 94.70% Conservative: 5
Best Local Similarity: 91.39% Mismatches: 1
Query Match: 91.11% Indels: 7
DB: 29 Gaps: 1
US-10-029-020-14_COPY_1250_1400 (1-151) x AY413477 (1-3190)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 1699 TACATCCGAGATCTCCCTCTGGAAATGTCCACCAATCTCTGGAGATG----- 1749
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrIleuAlaThrAspProMetSer 40
Db 1750 -----AGTCACAGCCACGACACAAATACTACTCTGGTACAGACCCCATGAGT 1797
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 1798 GGCGCGCGTCTCTCTCTGACACCAACAGCCGCGGGTCTTCAAGGTCAAGTCCACCA 1857
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 1858 GTGTGAAGACCTGTGTGAAGAACTCCGAGGTGTAGCAGGGACTGTGTGACCAAGTGCCTC 1917
QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 1918 CCCTTTGATGATACCTCGCTCGGAGATGTGGGAAGGCCACAGAACCACTACCTAAC 1977
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 1978 CCCAGGGGAATTACAGTGGACAAAGTTGGGCTCATTTATTTCGTGGACGCCACCATG 2037
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 2038 AGACGTGTGATCAAAATGGATCATCTCCACTTTGCTGGGCTCCATGACCTCACCTCG 2097

```

QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
DB 2098 GCCAGGCCCTCAGCTGTGACTCCGTCATGGAG 2130

CB524534 770 bp mRNA linear EST 09-JUL-2003
UI-M-FY0-cey-b-20-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:6843669 5', mRNA sequence.
ACCESSION CB524534
VERSION CB524534.1 GI:29357889
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
FEATURES
    Location/Qualifiers
        1..770
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /clone="IMAGE:6843669"
            /tissue_type="whole brain"
            /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
            /lab_host="DH10B (T1 phage resistant)"
            /clone_lib="NIH_BMAP_FY0"
            /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
            Site 2: Not I; The library was constructed according
            Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
            1996. Denatured RNA was size fractionated on a 1% agarose
            gel. First strand cDNA synthesis was primed with oligo-dT
            primer containing a Not I site. Double strand cDNA was
            size selected according to mRNA size fraction, ligated
            with EcoR I adaptor, digested with NotI and then cloned
            directionally into pYX-Asc vector. The library tag
            sequence located between the Not I site and the polyA tail
            is AGCAGACAG. This library was created for the University
            Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
            Developing Mouse Nervous System', supported by National
            Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
            program coordinator."
ORIGIN
Alignment Scores:
Pred. No.: 1.55e-70 Length: 770
Score: 645.50 Matches: 128
Percent Similarity: 93.66% Conservative: 5
Best Local Similarity: 90.14% Mismatches: 1
Query Match: 82.54% Indels: 8
DB: 14 Gaps: 1

US-10-029-020-14_COPY_1250_1400 (1-151) x CB524534 (1-770)

QY 11 ValThrAsnLeuLeu-GluLeuArgAsnLysAspPheArgHisSerProAlaHi 30
DB 1 GTCACCAACATCCTCTGGAGATG-----AGTCACAGCCCGCACCA 39

```

```

QY 30 sLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPheLeuSerAspSerAsnSe 50
DB 40 CAAATACTACTACCTGGCTACAGACCCCATGAGTGGGGCCGCTTCTGTCTGACCAACAG 99

QY 50 rArgArgValPheLysLysSerThrValValValLysAspLeuValLysAsnSerCl 70
DB 100 CCGCGGGTCTTCAAGGTCAAGTCCACCACAGTGGTGAAGGACCTGGTGAAGAACTCCGA 159

QY 70 uValValAlaGlyThrGlyAspGlnCysLeuProPheAspThrArgCysGlyAspGl 90
DB 160 GGTGTAGCAGGACTGCTGACCACTGCTCCCTTTGATGATACCCGCTGGGAGATGG 219

QY 90 yGlyLysAlaThrGluAlaThrLeuThrAsnProArgGlyLeThrValAspLysPheGl 110
DB 220 TGGGAAAGCCACAGAGCCAGCTCACTAACCCAGGGGAATTACAGTGGCAAGATTTGG 279

QY 110 yLeuLeuTyrPheValAspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIleSe 130
DB 280 GCTCATTTATTTCTGTGGACGGCACCATGATCAGACGTGTGATCAAAATGAATCATCTC 339

QY 130 rThrLeuLeuGlySerAsnAspLeuThrSerAlaArgProLeuSerCysAspSerValMe 150
DB 340 CACTTTGCTGGGCTCCATGACCTCACTCGCCAGGCCCTCAGCTGTGACTCGGTCAT 399

QY 150 tAsp 151
DB 400 GGAG 403

RESULT 4
CF538162 623 bp mRNA linear EST 12-SEP-2003
LOCUS UI-M-GIO-cho-h-21-0-UI.r1 NIH_BMAP_GIO Mus musculus cDNA clone
DEFINITION IMAGE:30536156 5', mRNA sequence.
ACCESSION CF538162
VERSION CF538162.1 GI:34590144
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
FEATURES
    Location/Qualifiers
        1..623
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /clone="IMAGE:30536156"
            /tissue_type="whole brain"
            /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
            /lab_host="DH10B (T1 phage resistant)"
            /clone_lib="NIH_BMAP_GIO"
            /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
            Site 2: Not I; The library was constructed according
            Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
            1996. Denatured RNA was size fractionated on a 1% agarose
            gel. First strand cDNA synthesis was primed with oligo-dT
            primer containing a Not I site. Double strand cDNA was
            size selected according to mRNA size fraction, ligated
            with EcoR I adaptor, digested with NotI and then cloned
            directionally into pYX-Asc vector. The library tag
            sequence located between the Not I site and the polyA tail
            is AGCAGACAG. This library was created for the University
            Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
            Developing Mouse Nervous System', supported by National
            Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
            program coordinator."
ORIGIN
Alignment Scores:
Pred. No.: 1.55e-70 Length: 623
Score: 645.50 Matches: 128
Percent Similarity: 93.66% Conservative: 5
Best Local Similarity: 90.14% Mismatches: 1
Query Match: 82.54% Indels: 8
DB: 14 Gaps: 1

```

size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACGGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 3,31e-67 Length: 623
Score: 618.00 Matches: 110
Percent Similarity: 90.67% Conservatives: 26
Best Local Similarity: 73.33% Mismatches: 14
Query Match: 79.03% Indels: 0
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x CF538162 (1-623)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 41 TAGCTCCGCGGATATTCGCTGGGAATGTGACAAGTGTTTTAGAAGCTAAGAAATAAA 100

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 101 GATTTTATAGATAGTAGACACCGCTACAGATATCTACTGCTACGGACCGACGTCACC 160

QY 41 GlyValAlaPheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrVal 60
Db 161 GGAGATTGTACGCTCTGTATATACACCGCGAGAAATCTATCGCGCGAATCACTCAGC 220

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 221 GGAGCCAAAGACCTGACTAAACCGTGAAGTGTGCGAGGACCGGGACAGTGCCTT 280

QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyValAlaThrGluAlaThrLeuThrAsn 100
Db 281 CCTTTTATGACGACCGCAGGTGTGGGATGGAGGAGGTGTGGAACACAGCTCATGAGT 340

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetile 120
Db 341 CCCAAGGAATGCAATCGATAAGACGGAATCTACTTTCTTGTATGGAACCATGATC 400

QY 121 ArgArgIleAspGlnAsnGlyIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 401 AGAAGGTTTCATCAAAATGGAATCATATCAACTCTCTGGGCTCCAAACGACCTCAGTCA 460

QY 141 AlaArgProLeuSerCysAspSerValMet 150
Db 461 GCTCGACCTTTAACCTGTGATATAGCATG 490

RESULT 5
CF744767
LOCUS
DEFINITION
UI-M-GVO-clr-f-21-0-UI.r1 NIH BMAP_GVO Mus musculus cDNA clone
IMAGE:3621740 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 640)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: gcapbs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouseefi.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pX-5.

FEATURES
Location/Qualifiers
Source

1..640
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30621740"
/tissue_type="whole brain"
/dev_stage="1.5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_GVO"
/note="Organ: Brain; Vector: pX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to BMAP, Lennan and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGTGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 2.68e-65 Length: 640
Score: 603.00 Matches: 108
Percent Similarity: 90.48% Conservatives: 25
Best Local Similarity: 73.47% Mismatches: 14
Query Match: 77.11% Indels: 0
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x CF744767 (1-640)

QY 4 ArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLysAspPheArg 23
Db 1 CGGATATTCGCTCGGATGTGCAAGTGTTTTAGAAGTAAAGATTTAGTA 60

QY 24 HisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyValAla 43
Db 61 CATAGTAGCAACCCAGCTCACAGATACTACTGGCTACGACCCAGTCCACCGAGATTG 120

QY 44 PheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrValValLys 63
Db 121 TAGCTCTGTATTAACCCCGCAGATCTATCGCCCGGAAATCACTACGGGAGCCAAA 180

QY 64 AspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeuProPheAsp 83
Db 181 GACCTGACTTAAACCCGCTGAAGTGTGGCAGCGGCGGGAACAGTGCCTTCCTTTGAC 240

QY 84 AspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsnProArgGly 103
Db 241 GAGGCCAGGTGTGGGATGGAGCAAGGCTGTGGAAGCAACGCTCATGATGCCAAGGA 300

QY 104 IleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetileArgGile 123
Db 301 ATGGCAATCATGAAGAACCGGACTGATCTACTTTGTGATGAACCATGATCAGAACGTT 360

QY 124 AspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSerAlaArgPro 143
Db 361 GATCAAAATGGAATCATATCACTCTCTGGGCTCCAAACGACCTCAGTCCAGCTCGACCT 420

| | | | |
|---|---|---|--------|
| Query Match: | 72.19% | Indels: | 7 |
| DB: | 29 | Gaps: | 1 |
| US-10-029-020-14_COPY_1250_1400 (1-151) x AY405421 (1-5069) | | | |
| QY | 1 | TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys | 20 |
| DB | 872 | TATGTGCGGCGGATATCCCTTCTGGAATGTAAACAGTGTCTTAACTA----- | 922 |
| QY | 21 | AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer | 40 |
| DB | 923 | -----AGCAGCAACCCAGCTCATAGATACTACCTTGCACCGGATCCAGTCACG | 970 |
| QY | 41 | GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal | 60 |
| DB | 971 | GGAGATTGTAGTTTCTGCACAAACACCCCGCAGAAATTTATCGCCCAAGTCATTACG | 1030 |
| QY | 61 | ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu | 80 |
| DB | 1031 | GGGGCAAAAGACTTGAATAAATGCAGAGTGTGCGGGGACAGGGGAGCAGTGCCTT | 1090 |
| QY | 81 | ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn | 100 |
| DB | 1091 | CGGTTTGACGAGCGAGATGTGGGATGGAGGAAAGCGCGTGGAAAGCCACACTCATG | 1150 |
| QY | 101 | ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle | 120 |
| DB | 1151 | CCCAAGGAATGGCAGTTGATAGAAATGATTAATCTACTTTGTGTGATGGAACCATG | 1210 |
| QY | 121 | ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer | 140 |
| DB | 1211 | AGGAAGTTGACCAAAATGGAATCATATCACTCTTCTGGGCTCTAACGATTGTACTTCA | 1270 |
| QY | 141 | AlaArgProLeuSerCysAspSerValMet | 150 |
| DB | 1271 | GCCAGACCTTTAACTTGTGACACCCAGCATG | 1300 |
| RESULT 8 | | | |
| AY405420 | 5094 bp | DNA | linear |
| LOCUS | Homo sapiens HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence, | | |
| DEFINITION | genomic survey sequence. | | |
| ACCESSION | AY405420 | GI:39761394 | |
| VERSION | AY405420.1 | GI:39761394 | |
| KEYWORDS | GSS. | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | |
| TITLE | Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios | | |
| JOURNAL | Science | 302 (5652), 1960-1963 (2003) | |
| PUBMED | 14671302 | | |
| AUTHORS | 2 (bases 1 to 5094) | | |
| REFERENCE | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (16-NOV-2003) | Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | |
| COMMENT | This sequence was made by sequencing genomic exons and ordering them based on alignment. | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..5094 | | |
| | /organism="Homo sapiens" | | |
| | /mol_type="genomic DNA" | | |
| | /db_xref="taxon:9606" | | |

| | | | |
|---|---|--|--------|
| Gene | <1..>5094 | | |
| /locus_tag="HCM2218" | | | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 4.24e-59 | | |
| Score: | 564.50 | | |
| Matches: | 104 | | |
| Percent Similarity: | 86.00% | | |
| Conservative: | 25 | | |
| Best Local Similarity: | 69.33% | | |
| Mismatches: | 14 | | |
| Query Match: | 72.19% | | |
| Indels: | 7 | | |
| Gaps: | 1 | | |
| US-10-029-020-14_COPY_1250_1400 (1-151) x AY405420 (1-5094) | | | |
| QY | 1 | TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys | 20 |
| DB | 872 | TATGTGCGGCGGATATCCCTTCTGGAATGTAAACAGTGTCTTAACTA----- | 922 |
| QY | 21 | AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer | 40 |
| DB | 923 | -----AGCAGCAACCCAGCTCATAGATACTACCTTGCACCGGATCCAGTCACG | 970 |
| QY | 41 | GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal | 60 |
| DB | 971 | GGAGATTGTAGTTTCTGCACAAACACCCCGCAGAAATTTATCGCCCAAGTCATTACG | 1030 |
| QY | 61 | ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu | 80 |
| DB | 1031 | GGGGCAAAAGACTTGAATAAATGCAGAGTGTGCGCGGACAGGGGAGCAGTGCCTT | 1090 |
| QY | 81 | ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn | 100 |
| DB | 1091 | CGGTTTGACGAGCGAGATGTGGGATGGAGGAAAGCGCGTGGAAAGCCACACTCATG | 1150 |
| QY | 101 | ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle | 120 |
| DB | 1151 | CCCAAGGAATGGCAGTTGATAGAAATGATTAATCTACTTTGTGTGATGGAACCATG | 1210 |
| QY | 121 | ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer | 140 |
| DB | 1211 | AGGAAGTTGACCAAAATGGAATCATATCACTCTTCTGGGCTCTAACGATTGTACTTCA | 1270 |
| QY | 141 | AlaArgProLeuSerCysAspSerValMet | 150 |
| DB | 1271 | GCCAGACCTTTAACTTGTGACACCCAGCATG | 1300 |
| RESULT 9 | | | |
| AY405422 | 5087 bp | DNA | linear |
| LOCUS | Mus musculus HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence, | | |
| DEFINITION | genomic survey sequence. | | |
| ACCESSION | AY405422 | GI:39761396 | |
| VERSION | AY405422.1 | GI:39761396 | |
| KEYWORDS | GSS. | | |
| SOURCE | Mus musculus (house mouse) | | |
| ORGANISM | Mus musculus | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| AUTHORS | 1 (bases 1 to 5087) | | |
| REFERENCE | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | |
| TITLE | Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios | | |
| JOURNAL | Science | 302 (5652), 1960-1963 (2003) | |
| PUBMED | 14671302 | | |
| AUTHORS | 2 (bases 1 to 5087) | | |
| REFERENCE | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | |
| TITLE | Direct Submission | | |

JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

COMMENT

FEATURES

source

Location/Qualifiers
1..5087
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
41..5087
/locus_tag="HCM2218"

gene

ORIGIN

Alignment Scores:
Pred. No.: 5,65e-59 Length: 5087
Score: 563.50 Matches: 103
Percent Similarity: 86.00% Conservativeness: 26
Best Local Similarity: 68.67% Mismatches: 14
Query Match: 72.06% Indels: 7
DB: 29 Gaps: 1

US-10-029-020-14_COPY_1250_1400 (1-151) x AY405422 (1-5087)

QY 1 TyrileArgArgilePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
DB 871 TACGTCGGCGGATATCCCGTCTGGGAATGTGACACAGTGTTTAGAACTA----- 921
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
DB 922 -----AGTAGCAACCCAGCTCACAGATACCTACCTGGTACGACCCAGTCACC 969
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrVal 60
DB 970 GGAGATTGTACGTCCTCTGATCTACATACCCGACAGATCTATCGGCCGGAATCACTCAG 1029
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
DB 1030 GGAGCCAAAGACCTGACTATAAAGCGTGAAGTGTGGCAGGACCGGGGAACAGTGCCTT 1089
QY 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
DB 1090 CCCTTTTACAGGCGCCAGTGTGGGAGTGGAGCAAGGCTGTGGAAGCAACGCTCATGAGT 1149
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
DB 1150 CCCAAGGAATGGCAATCGATACGAACGAGCTGATCTACTTTGTGTGGAACCATGATC 1209
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
DB 1210 AGAAGATTGATCAAAATGGAATCATATCAACTCTCTGGGCTCCACGACCTCACGTCA 1269
QY 141 AlaArgProLeuSerCysAspSerValMet 150
DB 1270 GCTCGACCTTTAACCTGTGATACGATG 1299

RESULT 10

LOCUS

DEFINITION

BU705631 778 bp mRNA linear EST 15-JUL-2003
UI-M-POO-cac-k-16-0-UI.r1 NIH BMAP_F00 Mus musculus cDNA clone
IMAGE:6409479 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 778)
NTH-MGC <http://mgs.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-i@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

source

1..778
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6409479"
/issue_type="whole brain"
/dev_stages="embryo 12.5dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH BMAP_F00"
/note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;
Site 2: Not I: The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TGAGAGACC. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 4.75e-58 Length: 778
Score: 546.50 Matches: 102
Percent Similarity: 84.00% Conservativeness: 24
Best Local Similarity: 68.00% Mismatches: 17
Query Match: 69.88% Indels: 7
DB: 13 Gaps: 1

US-10-029-020-14_COPY_1250_1400 (1-151) x BU705631 (1-778)

QY 1 TyrileArgArgilePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
DB 281 TACGTCGGCGGATATCCCGTCTGGGAATGTGACAACTGTTTAGAACTA----- 331
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
DB 332 -----AGTAGCAACCCAGCTCACAGATACCTACCTGGTACGACCCAGTCACC 379
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrVal 60
DB 380 GGAGATTGTACGTCCTCTGATCTACATACCCGACAGATCTATCGGCCGGAATCACTCAGC 439
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
DB 440 GGAGCCAAAGACCTGACTATAAAGCGTGAAGTGTGGCAGGACCGGGGAACAGTGCCTT 499
QY 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
DB 500 CCCTTTTACAGGCGCCAGTGTGGGAGTGGAGCAAGGCTGTGGAAGCAACGCTCATGAGT 559
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
DB 560 CTCANAGGAATGGCAATCGATACGAACGAGCTGATCTACTTTGTGTGGAACCATGATC 619
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140

JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

COMMENT

FEATURES

source

Location/Qualifiers
1..5087
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
41..5087
/locus_tag="HCM2218"

gene

ORIGIN

Alignment Scores:
Pred. No.: 5,65e-59 Length: 5087
Score: 563.50 Matches: 103
Percent Similarity: 86.00% Conservativeness: 26
Best Local Similarity: 68.67% Mismatches: 14
Query Match: 72.06% Indels: 7
DB: 29 Gaps: 1

US-10-029-020-14_COPY_1250_1400 (1-151) x AY405422 (1-5087)

QY 1 TyrileArgArgilePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
DB 871 TACGTCGGCGGATATCCCGTCTGGGAATGTGACACAGTGTTTAGAACTA----- 921
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
DB 922 -----AGTAGCAACCCAGCTCACAGATACCTACCTGGTACGACCCAGTCACC 969
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrVal 60
DB 970 GGAGATTGTACGTCCTCTGATCTACATACCCGACAGATCTATCGGCCGGAATCACTCAG 1029
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
DB 1030 GGAGCCAAAGACCTGACTATAAAGCGTGAAGTGTGGCAGGACCGGGGAACAGTGCCTT 1089
QY 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
DB 1090 CCCTTTTACAGGCGCCAGTGTGGGAGTGGAGCAAGGCTGTGGAAGCAACGCTCATGAGT 1149
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
DB 1150 CCCAAGGAATGGCAATCGATACGAACGAGCTGATCTACTTTGTGTGGAACCATGATC 1209
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
DB 1210 AGAAGATTGATCAAAATGGAATCATATCAACTCTCTGGGCTCCACGACCTCACGTCA 1269
QY 141 AlaArgProLeuSerCysAspSerValMet 150
DB 1270 GCTCGACCTTTAACCTGTGATACGATG 1299

RESULT 10

LOCUS

DEFINITION

BU705631 778 bp mRNA linear EST 15-JUL-2003
UI-M-POO-cac-k-16-0-UI.r1 NIH BMAP_F00 Mus musculus cDNA clone
IMAGE:6409479 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 778)
NTH-MGC <http://mgs.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-i@mail.nih.gov

primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 4,82e-52 Length: 801
 Score: 499.00 Matches: 90
 Percent Similarity: 87.30% Conservative: 20
 Best Local Similarity: 71.43% Mismatches: 16
 Query Match: 63.81% Indels: 0
 DB: 14 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x CD803099 (1-801)

QY 26 HisSerProLalHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPheLeu 45
 Db 2 AACAGCCAGGACACAAAGTATTACTTGGCTGTGGACCCGGTGACCGGCTCACTCTACGTC 61
 QY 46 SerAspSerAsnSerArgValPheLysLeuValPheLysSerThrValValValLysAspLeu 65
 Db 62 TCTGACACCAACAGTCCCGGAATCTACGAGTCAAGTCTCTGAGCGGACCCAAAGACCTG 121
 QY 66 ValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeuProPheAspThr 85
 Db 122 GCTGGAATTCGGAAGTTGTGGAGGAGCTGGCGAACAAATGCTACCTTTGATGAAGCC 181
 QY 86 ArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsnProArgGlyLeuThr 105
 Db 182 CGCTGTGGGATGAGGGAAGGCTGTGGACGCCACCTCATGAGCCCGGAGGTATTGCA 241
 QY 106 ValAspLysPheGlyLeuLeuTyrPheValAspGlyThrMetLleArgLleAspGln 125
 Db 242 GTAGACAAGAAATGGGCTATTGTATTTGTTGATGCCACCATGATCCGGAAGTGGACCA 301
 QY 126 AsnGlyLleLleSerThrLeuLeuGlySerAsnAspLeuThrSerAlaArgProLeuSer 145
 Db 302 AACGGATATCTCCACCTGCTGGGCTCCATGACCTCAGCTGTCCGACCACTGAGC 361
 QY 146 CysAspSerValMetAsp 151
 Db 362 TGTGACTCGAGCATGGAC 379

RESULT 13

CF530821 692 bp mRNA linear EST 12-SEP-2003
 LOCUS UI-M-GHO-cgu-g-06-0-UI.r1 NIH_BMAP_GHO Mus musculus cDNA clone
 DEFINITION IMAGE:30356597 5', mRNA sequence.

ACCESSION

VERSION CF530821

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 692)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

Unpublished (1999)

JOURNAL

Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgapbs-x@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at

<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

FEATURES

source

Location/Qualifiers
 1..692
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30356597"
 /tissue_type="Whole brain"
 /dev_stage="1, 5, and 15 days newborn"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_GHO"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CGAAGTGAAT. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 3.01e-50 Length: 692
 Score: 484.00 Matches: 91
 Percent Similarity: 85.71% Conservative: 23
 Best Local Similarity: 68.42% Mismatches: 17
 Query Match: 61.89% Indels: 2
 DB: 14 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x CF530821 (1-692)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
 Db 297 TAGCTCGGCGGATATTCGCTCTGGGAATGTCACAAAGTCTTTTAGAATAAGAAATAAA 356
 QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
 Db 357 GATTTAGACATAGTAGCAACCCAGCTCACAGATACCTACCTGGCTACGGACCCAGTCACC 416
 QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
 Db 417 GGAGATTGTACGTCCTCTGATACATAACACCCGAGAACTATCGCCGGAATCACTCACG 476
 QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
 Db 477 GGAGCCAAAGACCTGACTAAACCGTGAAGTGGTGGAGGACCGGGGACAGTGCCTT 536
 QY 81 -ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAs 100
 Db 537 CCCTTTTGACGAGCCAGGTGTGGGATGGAGCAAGGCTGTGGAAGCAACGCTCATGAG 596
 QY 100 nProArgGlyIleThrValAspLysPheGlyLeuIleTyr-PheValAspGlyThrMetI 120
 Db 597 TCNCANAGGAATGGCAATCGATAAGAACGAGCTGATCTACTTTTGTGTGATGGACCATGA 656
 QY 120 leArgArgIleAspGlnAsnGlyIleIleSerThr 131
 Db 657 TCAGAAAGGTTGATCANNATGGAATCATATCAACT 691

RESULT 14

CD352566

LOCUS

CD352566 724 bp mRNA linear EST 09-JUL-2003

DEFINITION UI-M-GIO-cgk-o-15-0-UI.r1 NIH BMAP_GIO Mus musculus cDNA clone
 IMAGE:6856648 5', mRNA sequence.
 CD352566
 CD352566.1 GI:31144077
 EST.
 Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 724)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES

source

Location/Qualifiers
 1..724
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone_lib="NIH BMAP_GIO"
 /tissue_type="whole brain"
 /dev_stage="embryo.13.5,14.5,16.5,17.5dpc"
 /lab_host="PH10B (T1 phage resistant)"
 /note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGCAGACAG. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). 'Gene discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 4.57e-47 Length: 724
 Score: 459.00 Matches: 83
 Percent Similarity: 88.89% Conservative: 21
 Best Local Similarity: 70.94% Mismatches: 13
 Query Match: 58.70% Indels: 0
 DB: 14 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x CD352566 (1-724)

QY 34 LeuAlaThrAspProMetSerGlyAlaValPheLeuSerAspSerAsnSerArgVal 53
 Db 2 CTGGCTACGGACCGACGAGATTGTGAGTCTCTGATCTACTACACCCGACGATC 61
 QY 54 PheIleuSerThrValValValValValValValValValValValValValVal 73
 Db 62 TATCGGCGGAATCACTACCTACGGGAGCCAAAGACCTGACTAAACACGCTGAAGTGTGCA 121
 QY 74 GlyThrGlyAspGlnCysLeuProPheAspThrArgCysGlyAspGlyGlyVal 93
 Db 122 GGGACCGGGGACAGTGCCTTCCCTTTGACGAGCCAGGTGTGGGGATGGAGGCAAGGCT 181

QY 94 ThrGluAlaThrLeuThrAsnProArgGlyIleThrValAspLysPheGlyLeuIleTyr 113
 Db 182 GTGGAGCAACGCTCATGAGTCCCAAGGAATGCAATCGATAAGACCGACTGATCTAC 241
 QY 114 PheValAspGlyThrMetIleArgGlyLeuAspGlnAsnGlyIleIleSerThrLeuLeu 133
 Db 242 TTTGTGATGGAACCATGATCAGAAAGGTTGATCAAAATGGAATCATATCAACTCTCTCG 301
 QY 134 GlySerAsnAspLeuThrSerAlaArgProLeuSerCysAspSerValMet 150
 Db 302 GGTCCACGACCTCAGCTCAGCTCGACCTTAACCTGTGATAGTAGCATG 352

RESULT 15

CD803315
 LOCUS CD803315
 DEFINITION UI-M-GVO-chu-m-02-0-UI.r1 NIH BMAP_GVO Mus musculus cDNA clone
 IMAGE:30545857 5', mRNA sequence.
 CD803315
 CD803315.1 GI:32462141
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE 1 (bases 1 to 789)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES

source

Location/Qualifiers
 1..789
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone_lib="NIH BMAP_GVO"
 /tissue_type="whole brain"
 /dev_stage="1,5, and 15 days newborn"
 /lab_host="PH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP_GVO"
 /note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated with
 EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CGAACTGAAT. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). 'Gene discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 5.51e-44 Length: 789
 Score: 435.00 Matches: 84
 Percent Similarity: 86.99% Conservative: 23
 Best Local Similarity: 68.29% Mismatches: 16

```
Query Match: 55.63% Indels: 2
DB: 14 Gaps: 0
US-10-029-020-14_COPY_1250_1400 (1-151) x CD803315 (1-789)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 398 TACGTCCGGCGGATATCCCGTCTGGGAATGTGACAAGTGTTTAGAACTAAGAATAAA 457
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 458 GATTTTAGACATAGTAGCAACCCAGCTCACGATACCTGGCTACGACCCAGTCACC 517
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 518 GGAGATTGTGACGCTCTGATACTAAACCCGCGAGATCTATCGCCGGAATCAGTCACG 577
QY 61 ValValLysAspLeuValLysAsnSerGluValValIleGlyThrGlyAspGlnCysLeu 80
Db 578 GGAGCCAAAGACCTGACTAAAAACGCTGAAGTGTGGCAGGACCGGGGACAGTGCCTT 637
QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 638 CCCITTGACGAGGCCAGGTGTGGGATGGAGGCAAGGCTGTGGAAGC-ACGCTCATGAGT 696
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 697 CNCAAGGAATGGCAATCGATAGAACGGAATGATCTACTTTGTGTGGACC-ATGATC 755
QY 121 ArgArgIle 123
Db 756 AGAAGGTTG 764
```

Search completed: August 14, 2004, 18:03:23
Job time : 1784.9 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 256.137 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_1250_1400

Perfect score: 782

Sequence: 1 YIRIFPSGVNTILRNK.....LLGSNDLSARPLSCDSVMD 151

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020.@CNG 1.1.1968 @runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 782 | 100.0 | 8354 | 6 | Abs52100 Human TEN |
| 2 | 759.5 | 97.1 | 8438 | 6 | Abn85378 Human NOV |
| 3 | 727.5 | 93.0 | 8645 | 6 | Abn78652 Human CDN |
| 4 | 619 | 79.2 | 8473 | 6 | Abq82345 Human NOV |
| 5 | 619 | 79.2 | 8487 | 6 | Abq82346 Human NOV |
| 6 | 619 | 79.2 | 8645 | 6 | Abq82344 Human NOV |
| 7 | 619 | 79.2 | 8675 | 6 | Abq82343 Human NOV |
| 8 | 613 | 78.4 | 9058 | 7 | Acc72051 BCU0205A |

| | | | | | | |
|----|-------|------|--------|---|-------------|---------------------|
| 9 | 613 | 78.4 | 9695 | 7 | ACC72052 | Acc72052 BCU0205B |
| 10 | 613 | 78.4 | 9729 | 5 | AAS14089 | Aas14089 Human FCT |
| 11 | 613 | 78.4 | 9729 | 9 | ADB32028 | Adb32028 Human FCT |
| 12 | 613 | 78.4 | 9826 | 5 | AAS14085 | Aas14085 Human FCT |
| 13 | 613 | 78.4 | 9826 | 9 | ADB32023 | Adb32023 Human FCT |
| 14 | 571.5 | 73.1 | 4245 | 7 | AAI60066 | Aai60066 Human Pco |
| 15 | 534.5 | 68.4 | 12879 | 6 | ABK92230 | Abk92230 Prostate |
| 16 | 534.5 | 68.4 | 13202 | 4 | AAK51828 | Aak51828 Human pol |
| 17 | 281.5 | 36.0 | 10242 | 4 | ABL29075 | Abi29075 Drosophil |
| 18 | 281.5 | 36.0 | 17131 | 4 | ABL29074 | Abi29074 Drosophil |
| 19 | 253 | 32.4 | 1764 | 4 | ABL04855 | Abi04855 Drosophil |
| 20 | 253 | 32.4 | 4925 | 4 | ABL04854 | Abi04854 Drosophil |
| 21 | 101 | 12.9 | 1482 | 6 | ABZ14144 | Abz14144 Arabidops |
| 22 | 91 | 11.6 | 247 | 3 | AAI18884 | Aai18884 Human sec |
| 23 | 90.5 | 11.6 | 110000 | 4 | AAI19682_10 | Continuation (11 o |
| 24 | 90.5 | 11.6 | 110000 | 4 | AAI19683_10 | Continuation (11 o |
| 25 | 87.5 | 11.2 | 1842 | 3 | AAC39026 | Aac39026 Arabidops |
| 26 | 85 | 10.9 | 2550 | 9 | ADB61918 | Adb61918 Human CDN |
| 27 | 85 | 10.9 | 2700 | 6 | ABN59835 | Abn59835 Novel hum |
| 28 | 85 | 10.9 | 4308 | 6 | ABS67814 | Abs67814 Human rec |
| 29 | 84 | 10.7 | 6437 | 4 | AAK82889 | Aak82889 Human imm |
| 30 | 83 | 10.6 | 110000 | 4 | AAI19682_39 | Continuation (40 o |
| 31 | 83 | 10.6 | 110000 | 4 | AAI19683_39 | Continuation (40 o |
| 32 | 82 | 10.5 | 4377 | 4 | RAA55933 | RAA55933 Streptoco |
| 33 | 81 | 10.4 | 4389 | 7 | ABX05900 | Abx05900 S. pneumo |
| 34 | 81 | 10.4 | 9607 | 2 | AAV52213 | Aav52213 Streptoco |
| 35 | 81 | 10.4 | 15438 | 7 | ABT42143 | Abt42143 Toxicity |
| 36 | 81 | 10.4 | 110000 | 7 | ABS56454_01 | Continuation (2 of |
| 37 | 79.5 | 10.2 | 2736 | 2 | AAK87478 | Aax87478 Rat vanil |
| 38 | 79.5 | 10.2 | 2736 | 2 | AAK19730 | Aax19730 Rat VRPP- |
| 39 | 79 | 10.1 | 1610 | 2 | AAI13498 | Aaxi13498 Enterococ |
| 40 | 79 | 10.1 | 1610 | 6 | ABS99293 | AbS99293 Enterococ |
| 41 | 78.5 | 10.0 | 14042 | 2 | AAT28774 | Aat28774 Human pla |
| 42 | 78.5 | 10.0 | 14042 | 2 | AAV05995 | Aav05995 Human pla |
| 43 | 78.5 | 10.0 | 14044 | 2 | AAT28776 | Aat28776 Human par |
| 44 | 78.5 | 10.0 | 14044 | 2 | AAV05997 | Aav05997 Human par |
| 45 | 78.5 | 10.0 | 14080 | 2 | AAT28775 | Aat28775 Human kid |

ALIGNMENTS

RESULT 1
ABS52100
ID ABS52100 standard; DNA; 8354 BP.
XX
AC ABS52100;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human TEN-M4-like gene.
XX

Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
cell signal processing; metabolic pathway modulation; metabolic disorder;
obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
memory defect; infertility; congenital heart defect; hair growth;
pigmentation disorder; endocrine disorder; respiratory disease; health;
gastro-intestinal disease; reproductive; neurological disease;
bone marrow transplantation; endocrine disease; allergy; inflammation;
neurological disorder; urinary system disorder; age-related disorder;
neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
adipocyte complement-related C1q tumour necrosis factor; out at first;
beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
type 1a membrane sushi-containing domain; butyrophilin;
type 1a membrane sushi-containing domain containing; SNP; gene; ds;
single nucleotide polymorphism.

Homo sapiens.

Key Location/Qualifiers
variation replace(117,G)
/*tag= a

FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN WO200257453-A2.
XX
XX 25-JUL-2002.
XX
XX 19-DEC-2001; 2001WO-US050331.
XX
XX 19-DEC-2000; 2000US-0256704P.
XX 20-DEC-2000; 2000US-0257314P.
XX 02-MAY-2001; 2001US-0286153P.
XX 29-MAY-2001; 2001US-0294075P.
XX 24-JUL-2001; 2001US-0307506P.
XX 10-AUG-2001; 2001US-0311590P.
XX 10-AUG-2001; 2001US-0311613P.
XX 29-AUG-2001; 2001US-0315617P.
XX 14-SEP-2001; 2001US-0322359P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
XX Stone DJ, Anderson D, Shinkets RA, Burgess CE, Zerhusen BD, Liu X;
XX Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX WPI: 2002-590744/53.
XX P-PSDB; ABG70388.

XX
XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
XX atherosclerosis, metabolic disorders, diabetes, obesity, infectious
XX disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
XX cancer.

XX Claim 8; Page 50-52; 318pp; English.

XX The present invention relates to new NOVX polypeptides. The invention is
XX useful for treating or preventing a NOVX-associated disorder such as
XX cardiomyopathy or atherosclerosis, where the disorder is related to cell
XX signal processing and metabolic pathway modulation in a subject.
XX preferably human. The invention is also useful for treating metabolic
XX disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
XX disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
XX disorders, haematopoietic disorders and various cancers. The molecules of
XX the invention are also useful for treating or preventing cirrhosis,
XX pancreatitis, learning and memory defects, infertility, congenital heart
XX defects, acne, hair growth, pigmentation disorders, reproductive disorders,
XX respiratory disease, gastro-intestinal diseases, endocrine diseases,
XX neurological diseases, bone marrow transplantation, endocrine diseases,
XX allergy and inflammation, nephrological disorders, urinary system
XX disorders, neuropsychiatric disorders and age-related disorders. The
XX present nucleic acid sequence represents a NOVX gene. This sequence
XX encodes a NOVX protein of the invention

XX Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;

XX Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 7,99e-95 | Length: | 8354 |
| Score: | 782.00 | Matches: | 151 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-029-020-14_COPY_1250_1400 (1-151) x ABS52100 (1-8354)

Qy 1 TyrilleArgGlyPheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3782 TACATTAGAGGATCTTCCCTCTGGAAATGTCCACCAACATCTTAGAGCTAGAGATAAA 3841

Qy 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3842 GATTTTCAGACATAGTCCAGTCCAGCACACAAATACTACTTGGCCACAGACCCCATGAGT 3901
Qy 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3902 GGGGCGGTCTTCTTCTGACAGCAACAGCGCGGGTCTTTAAATCAAGTCCACTGTG 3961
Qy 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3962 GTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTGGGACAGGACACCGTGGCTTC 4021
Qy 81 ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4022 CCCTTTGATGACACTCGCTCGGGGATGGTGGGAAGGCCACAGAGCCACACTCACC 4081
Qy 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4082 CCCAGGGGTATTACAGTGGACAAAGTTGGGCTGATCTACTTCGTGGTGGCCACTGATC 4141
Qy 121 ArgAlaGlyLeuAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4142 AGACGCATCGATCAGATGGGATCATCTCCACCTGCTCGGCTCTTAATGATCTCATCA 4201
Qy 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4202 GCCCGGCCACTCAGCTGTGATCTCTCATGGAT 4234

RESULT 2

ABN85378

ID ABN85378 standard; DNA; 8438 BP.

XX AC ABN85378;

XX 21-OCT-2002 (first entry)

XX Human NOV1, TEN-M4 like protein, coding sequence.

XX Human; NOV1: cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
XX Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
XX Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
XX Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
XX Gene Therapy; NOV; cancer; heart disease; inflammation;
XX autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
XX asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
XX infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
XX TEN-M4 like protein; chromosome 11; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 4..8395

XX /tag= a

XX /trans_except= (pos: 1138..1147,aa:Met)

XX /product= "NOV1 protein"

XX WO200255704-A2.

XX 18-JUL-2002.

XX 09-JAN-2002; 2002WO-US000554.

XX 09-JAN-2001; 2001US-0260417P.

XX 10-JAN-2001; 2001US-0260831P.

XX 28-FEB-2001; 2001US-0272338P.

XX 09-MAR-2001; 2001US-0274976P.

XX 18-APR-2001; 2001US-0284704P.

XX (CURA-) CURAGEN CORP.

XX Padigar M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;

XX Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;

PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
 PI Alsbrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
 PI Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
 PI Gunther E, Stone DJ;
 DR WPI; 2002-590674/63.
 DR P-PSDB; ABB98401.
 XX NOVX polypeptides and encoding polynucleotides, useful for preventing or
 PT treating NOVX-associated disorders e.g. cancer, inflammation, or
 PT Alzheimer's disease, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Claim 9; Page 8-9; 358pp; English.
 XX
 CC The present sequence is a coding sequence for a NOV protein. The NOV
 CC proteins and coding sequences are useful for treating or preventing NOV-
 CC associated disorders or in the manufacture of a medicament for treating
 CC the disorders, such as cancer, heart disease, inflammation, autoimmune
 CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
 CC Iga nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
 CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
 CC and other wasting disorders associated with chronic diseases. NOV1 is a
 CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
 XX
 SQ Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 9.38e-92 Length: 8438
 Score: 759.50 Matches: 151
 Percent Similarity: 96.79% Conservative: 0
 Best Local Similarity: 96.79% Mismatches: 0
 Query Match: 97.12% Indels: 5
 DB: 6 Gaps: 2

US-10-029-020-14_COPY_1250_1400 (1-151) x ABB85378 (1-8438)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeu-----Arg 18
 Db 3764 TACATTAGAGGATCTCCCTCTGGNAATGTACCACATCTCTAGAGCTCAGGAGGTGAGA 3823

QY 19 AsnLysAspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPro 38
 Db 3824 AATAAGATTTCAGACATAGTCACAGTCCAGCACAAATACTACTCTGGCCACAGACCCC 3883

QY 39 MetSerGlyAlaValPheLeuSerAspSerAsnSerArgValPheLysIleLysSer 58
 Db 3884 ATGAGTGGGCGGCTCTCTCTCTGACAGCAACAGCGCGGGTCTTTAAATCAAGTCC 3943

QY 59 ThrValValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGln 78
 Db 3944 ACTGTGTGTGTGAAGACCTTGTCAAGAACTCTGAGGTGTGTCGGGGACAGGTGACCAG 4003

QY 79 CysLeuProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeu 98
 Db 4004 TGCTCTCCCTTTGATGACACTCGCTCGGGGATGTGGGAAGCCACAGAGCCACATC 4063

QY 99 ThrAsnProArg-----GlyIleThrValAspLysPheGlyLeuIleTyrPheVal 115
 Db 4064 ACCAATCCCGGGGTCCTCCCGGACATTTACAGTGGACAAGTTTGGGCTGATCTACTTCGTG 4123

QY 116 AspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySer 135
 Db 4124 GATGGCACCATGATCAGCGCATCGATCAGATGGGATCATCTCCACCTTCGCTCGGCTCT 4183

QY 136 AsnAspLeuThrSerAlaArgProLeuSerCysAspSerValMetAsp 151
 Db 4184 AATGATCTCACATCAGCCCGGCACTCAGCTGTGATTCTGTCTCATGAT 4231

RESULT 3
 ABS78652
 ID ABS78652 standard; cDNA; 8645 BP.
 XX

AC ABS78652;
 XX 16-DEC-2002 (first entry)
 XX Human cDNA encoding CGDD10, INCYTE 7488573CB1.
 DE XX
 XX Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
 KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
 KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
 KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
 KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
 KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
 KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
 KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
 KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
 KW multiple sclerosis; glomerulonephritis; irritable bowel syndrome;
 KW rheumatoid arthritis; osteoarthritis; osteoporosis; pancreatitis;
 XX
 OS Homo sapiens.
 XX WO200272830-A2.
 XX 19-SEP-2002.
 XX
 XX 08-FEB-2002; 2002WO-US003715.
 XX
 XX 09-FEB-2001; 2001US-0268111P.
 XX 23-FEB-2001; 2001US-0271175P.
 XX 08-MAR-2001; 2001US-0274503P.
 XX 09-MAR-2001; 2001US-0274552P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
 PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
 PI Lu DM, Richardson TW, Tran UK, Khare R, Walla NK;
 DR WPI; 2002-723356/78.
 DR P-PSDB; ABG97359.
 XX
 XX New human proteins associated with cell growth, differentiation and
 PT death, useful for diagnosing, treating or preventing autoimmune or
 PT inflammatory disorders (e.g. AIDS, allergy or anaemia), cancer,
 XX atherosclerosis or hepatitis.
 XX
 XX Claim 5; Page 175-178; 181pp; English.
 XX
 CC The invention relates to an isolated polypeptide comprising CGDD1-12
 CC (cell growth, differentiation and death), a naturally occurring amino
 CC acid sequence at least 90% identical to CGDD, a biologically active
 CC fragment or an immunogenic fragment. Also included are the
 CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
 CC comprising a promoter sequence operably linked to the CGDD
 CC polynucleotide, a cell transformed with the recombinant polynucleotide,
 CC a transgenic organism comprising the recombinant polynucleotide, an anti-
 CC body/antagonist of CGDD or alter the expression of CGDD polynucleotide and a
 CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
 CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
 CC agonists and antagonists are useful for diagnosing, treating or
 CC preventing disorders associated with aberrant expression of CGDD,
 CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
 CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
 CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
 CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
 CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
 CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
 CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
 CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
 CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
 CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
 CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
 CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic

CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein

XX SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 2.21e-87 Length: 8645
Score: 727.50 Matches: 144
Percent Similarity: 95.36% Conservative: 0
Best Local Similarity: 95.36% Mismatches: 7
Query Match: 23.03% Indels: 1
DB: 6 Gaps:

US-10-029-020-14_COPY_1250_1400 (1-151) x ABS78652 (1-8645)

Qy 1 TyrileArgAlaGilePheProSerGlyAsnValThrAsnileLeuGluLeuArgAsnlys 20
Db 3852 TACATTAGAGGAGTCTCCCTCTGGAATGTCCACCAATCTAGAGCTG----- 3902
Qy 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3903 -----AGTCACAGTCCAGCACACAATACTACTCTGCCACAGACCCCATGAGT 3950
Qy 41 GlyAlaValPheLeuSerAspSerAsnSerArgGValPheLysIleLysSerThrVal 60
Db 3951 GGGGGCGTCTCTCTTCTGACAGCAACAGCCGGCGGCTTTTAAATCAAGTCCACTGTG 4010
Qy 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 4011 GTGGTGAGGACCTTGTCAAGACTCTGAGTGTTGTGGGGACAGGTGACCATGCTTC 4070
Qy 81 PropheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4071 CCCTTTGATGACACTCGCTGCGGGATGTTGGGAAGCCACAGACCACTACCCAT 4130
Qy 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetile 120
Db 4131 CCCAGGGGCAATACAGTGACAAAGTTGGGCTGATCTCTTGTGGATGGACCATGATC 4190
Qy 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4191 AGACGATCGATCAGATGGATCATCTCCACCTGCTGGCTCTAATGATCTCATCATCA 4250
Qy 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4251 GCCCGGCCACTCAGCTGTGATTTCTGTATGGAT 4283

RESULT 4

ID ABQ82345 standard; cDNA; 8473 BP.

XX AC ABQ82345;

XX 17-DEC-2002 (first entry)

XX Human NOV15c encoding cDNA SEQ ID NO:39.

XX Human; NOVX; cytosolic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nontropic; antidiabetic; antiinflammatory; fungicide;
KW antineuritic; antiarthritic; immunosuppressive; antiinfective; virucide;
KW antineuritic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; infection;
KW ulcerative colitis; allergic reaction; autoimmune haemolytic anaemia;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

OS

FH Key Location/Qualifiers
FT CDS 258..8144
FT /*tag= a
FT /product= "NOV15c"

XX WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

XX 29-DEC-2000; 2000US-0258928P.

XX 02-JAN-2001; 2001US-0259415P.

XX 04-JAN-2001; 2001US-0259785P.

XX 20-FEB-2001; 2001US-0269814P.

XX 09-MAR-2001; 2001US-0279863P.

XX 29-MAR-2001; 2001US-0279832P.

XX 29-MAR-2001; 2001US-0279833P.

XX 13-APR-2001; 2001US-0283889P.

XX 18-APR-2001; 2001US-0284447P.

XX 25-APR-2001; 2001US-0286683P.

XX 29-MAY-2001; 2001US-0294080P.

XX 16-AUG-2001; 2001US-0312915P.

XX 17-AUG-2001; 2001US-0313325P.

XX 17-SEP-2001; 2001US-0322699P.

XX 26-NOV-2001; 2001US-0333350P.

XX (CURA-) CURAGEN CORP.

XX Spteck KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;

XX Malyankar U, Shamkets RA, Tchernev VT, Spaderna SK, Gorman L;

XX Kiyuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;

XX Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;

XX Gunther E, Smithson G, Millet I, Macdougall JR;

XX WPI; 2002-732706/79.

XX P-PSDB; ABP53588.

XX New NOVX polypeptides and polynucleotides useful for treating NOVX-

XX associated disorders, such as cancers, neurological disorders, disorders

XX of vesicular transport, gastrointestinal disorders, and autoimmune

XX diseases.

XX Claim 8; Page 119-121; 444pp; English.

XX The present invention describes novel human proteins designated NOVX,

XX where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,

XX cytosolic, anticonvulsant, cerebroprotective, nontropic, cardiovascular,

XX antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,

XX immunosuppressive, antiallergic, antineuritic, antibacterial, fungicide,

XX protozoacide and antihelminthic activities, and can be used in gene

XX therapy. The NOVX proteins, nucleotides or antibodies can be used in the

XX manufacture of a medicament for treating a syndrome associated with a

XX human disease selected from NOVX-associated disorder, such as cancers

XX (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,

XX ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,

XX ischaemic cerebrovascular disease, Alzheimer's disease or Pick's

XX disease), disorders of vesicular transport (e.g. cystic fibrosis,

XX diabetes mellitus, Grave's disease, or goitre), gastrointestinal

XX disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),

XX autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic

XX anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic

XX and protozoal infections. The NOVX proteins can be used as immunogens to

XX produce antibodies and as vaccines. The NOVX nucleotide sequences may be

XX used in chromosome mapping, identifying individuals from minute

XX biological samples (tissue typing), and in forensic identification of a

XX biological sample. The present sequence encodes human NOV15c, which is

XX located on chromosome 4

XX SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 1.27e-72

Length:

8473

```
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x ABQ82345 (1-8473)

QY 1 TyrlleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3600 TACGTGCGCGGATATCCCTTCTCGAATGTAACAAAGTGTCTTAAAGCTAAGAAATAAA 3659

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrlleValLeuAlaThrAspProMetSer 40
Db 3660 GATTTTAGACATAGCAGCAACCCAGCTCATAGATCTACTTGTGCAACGATCCAGTCCAG 3719

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLeuSerThrVal 60
Db 3720 GGAGATCTGTAGCTTCTGACACAAACACCCGCGAGATTTATCGCCCAAGTCACTTACG 3779

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3780 GGGGCAAAAGACTTGCATTAATAATGCAGAGTGTGCGAGGACAGGCGGACATGCTT 3839

QY 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 3840 CCGTTTGACGAGCGAGATGTGGGATGGAGGAGGCGGTGGAGCCACACTCATGAGT 3899

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrlleValAspGlyThrMetIle 120
Db 3900 CCCAAAGGAATGGCAGTGTGTAAGAATGATTAATCTACTTTGTGATGAACCATGATT 3959

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 3960 AGGAAGTGTGCAAAATGGGAATCATATCAACTTTCTGGGCTCTAAACGATTGACTTCA 4019

QY 141 AlaArgProLeuSerCysAspSerValMet 150
Db 4020 GCCAGACCTTTAACTTGTGACACCGACATG 4049

RESULT 5
ABQ82346
ID ABQ82346 standard; cDNA; 8487 BP.
XX
AC ABQ82346;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15d encoding cDNA SEQ ID NO:41.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; infection;
KW ulcerative colitis; gastric disorder; duodenal disorder; vaccine;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 299..8140
FT /*tag= a
FT /product= "NOV15d"
XX
PN W0200262999-A2.
XX
PD 15-AUG-2002.
XX
```

```
PF 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0259814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
```

SPYTEK KA, LI L, WOLENC AR, VERNET CAM, EISEN A, LIU X;
MALYANKAR U, SHIMKETS RA, TCHERNEV VT, SPADERNA SK, GORMAN L;
KEKUDA R, PATTURAJAN M, GUSEV V, GANGOLLI EA, GUO X, SHENOY S;
RASTELLI L, CASMAN SJ, BOLDOG F, BURGESS CE, EDINGER S, ELLERMAN K;
GUNTHER E, SMITHSON G, MILLET I, MACDOUGALL JR;

WPI; 2002-732706/79.
P-PSDB; ABP53589.

New NOVX polypeptides and polynucleotides useful for treating NOVX-associated disorders, such as cancers, neurological disorders, disorders of vesicular transport, gastrointestinal disorders, and autoimmune diseases.

Claim 8; Page 123-125; 444pp; English.

The present invention describes novel human proteins designated NOVX, where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective, cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular, antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide, immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide, protozoacide and antihelminthic activities, and can be used in gene therapy. The NOVX proteins, nucleotides or antibodies can be used in the manufacture of a medicament for treating a syndrome associated with a human disease selected from NOVX-associated disorder, such as cancers (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle, ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke, ischaemic cerebrovascular disease, Alzheimer's disease or Pick's disease), disorders of vesicular transport (e.g. cystic fibrosis, diabetes mellitus, Grave's disease, or goitre), gastrointestinal disorders (e.g. ulcerative colitis, or gastric and duodenal disorders), autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic and protozoal infections. The NOVX proteins can be used as immunogens to produce antibodies and as vaccines. The NOVX nucleotide sequences may be used in chromosome mapping, identifying individuals from minute biological samples (tissue typing), and in forensic identification of a biological sample. The present sequence encodes human NOV15d, which is located on chromosome 4.

Sequence 8487 BP; 2306 A; 2029 C; 2148 G; 2001 T; 0 U; 3 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 1,27e-72 | Length: | 8487 |
| Score: | 619.00 | Matches: | 111 |
| Percent Similarity: | 90.67% | Conservative: | 25 |
| Best Local Similarity: | 74.00% | Mismatches: | 14 |
| Query Match: | 79.16% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-029-020-14_COPY_1250_1400 (1-151) x ABQ82346 (1-8487)

QY 1 TyrlleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20

Db 3596 TACGTGGCGGATATTCCTCTTGGAAATGTACAACTGCTTAGAATAA 3655
QY 21 AspPheArgHisSerHisSerProAlaHisIysTyrLeuAlaThrAspProMetSer 40
Db 3656 GATTTTACATAGCAGCAGCAGCAGCTCATAGATACCTTGTCAACGATCCAGTCACG 3715
QY 41 GlyAlaValPheLeuSerAspSerHisSerArgValPheLeuIleIysSerThrVal 60
Db 3716 GGAGATCTGTAGTTTCTGACAAACACCCGACAGATTTATCGCCCAAGTCACTTACG 3775
QY 61 ValValIysAspLeuValIysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3776 GGGGCAAAAGACTTACTTAAATAATCGAAGTCGTGCGAGGACAGGGGAGCAATGCCCT 3835
QY 81 ProPheAspAspThrArgCysGlyAspGlyIysAlaThrGluAlaThrLeuThrAsn 100
Db 3836 CCGTTTGACGAGCGAGATGTGGGATGCGAGGAGCGCGTGGAAAGCCACACTCATGAT 3895
QY 101 ProArgGlyIleThrValAspIysPheGlyIleuIleTyrPheValAspGlyThrMetIle 120
Db 3896 CCCAAAGGAATGGCAGTGTATGAAGATGGAATTAATCTACTTTGTGATGAACCATGAT 3955
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 3956 AGGAAGTGTGACAAATAAGATCATATCACTCTCTGGGCTTAACGATTTGACTTCA 4015
QY 141 AlaArgProLeuSerCysAspSerValMet 150
Db 4016 GCCAGACCTTTAACTTGTGACACCCAGCATG 4045
RESULT 6
ABQ82344
ID ABQ82344 standard; cDNA; 8645 BP.
AC ABQ82344;
XX
DT 17-DEC-2002 (first entry)
DE Human NOV15b encoding cDNA SEQ ID NO:37.
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 151..8316
FT /tag= a
FT /product= "NOV15b"
XX
PN W0200262999-A2.
XX
PD 15-AUG-2002.
XX
FF 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0259814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR

PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313225P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shinkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
XX WPI; 2002-732706/79.
XX P-PSDB; ABP53587.
XX
XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
XX associated disorders, such as cancers, neurological disorders, disorders
XX of vesicular transport, gastrointestinal disorders, and autoimmune
XX diseases.
XX
XX Claim 8; Page 114-117; 444pp; English.
XX
XX The present invention describes novel human proteins designated NOVX,
XX where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
XX cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
XX antidiabetic, antiinflammatory, antianaemic, antibacterial, virucide,
XX immunosuppressive, antiallergic, antihelminthic activities, and can be used in gene
XX therapy. The NOVX proteins, nucleotides or antibodies can be used in the
XX manufacture of a medicament for treating a syndrome associated with a
XX human disease selected from NOVX-associated disorder, such as cancers
XX (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
XX ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
XX ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
XX disease), disorders of vesicular transport (e.g. cystic fibrosis,
XX diabetes mellitus, Grave's disease, or goitre), gastrointestinal
XX disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
XX autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
XX anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
XX and protozoal infections. The NOVX proteins can be used as immunogens to
XX produce antibodies and as vaccines. The NOVX nucleotide sequences may be
XX used in chromosome mapping, identifying individuals from minute
XX biological samples (tissue typing), and in forensic identification of a
XX biological sample. The present sequence encodes human NOV15b, which is
XX located on chromosome 4
XX
XX Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;
SQ
Alignment Scores:
Pred. No.: 1,31e-72 Length: 8645
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: Gaps: 6
US-10-029-020-14_COPY_1250_1400 (1-151) x ABQ82344 (1-8645)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3772 TACGTGGCGGATATTCCTCTTGGAAATGTACAACTGCTTAGAATAA 3831
QY 21 AspPheArgHisSerHisSerProAlaHisIysTyrTyrLeuAlaThrAspProMetSer 40
Db 3832 GATTTTACATAGCAGCAGCAGCTCATAGATACCTTGTCAACGATCCAGTCACG 3891
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgValPheIysIleLysSerThrVal 60

Db 4024 CCGTTTGACGAGCGCAGATGTGGGATGGAGGCGGTGGAAGCCACACTCATGACT 4083
 QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
 Db 4084 CCCAAGGAATGGCAGTTTGATAAGAATGGATTAAATCTACTTTGTTGATGAACCATGATT 4143
 QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 Db 4144 AGGAAGTTGACCAAAATGGAATCATATCACTCTCTGGGCTCTTAACGATTTGACTTCA 4203
 QY 141 AlaArgProLeuSerCysAspSerValMet 150
 Db 4204 GCCAGACCTTTAACTTGTGACACCGCATG 4233
 RESULT 8
 ID ACC72051 standard; DNA; 9058 BP.
 XX
 AC ACC72051;
 DT 08-JUL-2003 (first entry)
 XX
 DE BCU0205A gene #SEQ ID 79.
 XX
 KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
 KW drug discovery; clinical medicine; forensic medicine; gene;
 KW chromosome 5q33.3; ds.
 XX
 OS Homo sapiens.
 XX
 EN WO2003029421-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-US031287.
 XX
 PR 03-OCT-2001; 2001US-0326526P.
 PR 14-MAY-2002; 2002US-00144194.
 XX
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX
 PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
 XX
 DR WPI; 2003-381623/36.
 DR P-PSDB; ABR58317.
 XX
 PS New isolated human differentially-regulated breast cancer polynucleotide
 PT and polypeptide, useful for diagnosing, staging, prognosticating,
 PT preventing and/or treating diseases and conditions relating to breast
 PT cancer.
 XX
 PS Claim 2; SEQ ID NO 79; 127bp + Sequence Listing; English.
 XX
 CC The invention relates to isolated polynucleotides which are
 CC differentially-regulated in breast cancer. The methods and compositions
 CC of the present invention are useful for detecting, diagnosing, staging,
 CC monitoring, prognosticating, preventing and/or treating diseases and
 CC conditions relating to breast cancer, and may be used in gene therapy or
 CC antisense therapy. They can also be used in research, drug discovery,
 CC clinical medicine and forensic medicine. Sequences given in records
 CC ACC72012-ACC72074 represent polynucleotides of the invention that are
 CC differentially-regulated in breast cancer. NOTE: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC fcp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;

Query Match: 78.33% Indels: 0
 DB: 7 Gaps: 0
 US-10-029-020-14_COPY_1250_1400 (1-151) x ACC72051 (1-9058)
 QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuLeuArgAsnLys 20
 Db 3146 TACATCCGACGCATCTTTCCCTCTCGAATGTGACACAGCATCTTGGAGTTACGAAATAA 3205
 QY 21 AspPheArgHisSerHisSerProAlaHisIleTyrTyrLeuAlaThrAspProMetSer 40
 Db 3206 GAGTTTAAACATAGCAACACCCAGCACACAAGTACTACTTGGCAGTGGACCCCGTGTC 3265
 QY 41 GlyValaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
 Db 3266 GGCTCGCTCTACGTGTCGACACCAACACAGCAGAGAATCTACCGCGTCAAGTCTCTGAGT 3325
 QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
 Db 3326 GGAACCAAAAGACCTGGCTGGGAATTGCGAAGTTGTGGCAGGACGGGAGCGAGTGTCTA 3385
 QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
 Db 3386 CCTTTGATGAAGCCCGCTGCGGGATGGAGGAGGCCATAGATGCAACCTGATGAGC 3445
 QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
 Db 3446 CCGAGAGGTATTCCAGTAGACAAGATGGCTCATCTTTTGTGATGCCACCATGATC 3505
 QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 Db 3506 CGGAGGTTGACCAAGATGGAATCATCTCACCTCTGGGCTCCAATGACCTCACTGTC 3565
 QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
 Db 3566 GTCCGCGCGCTGAGCTGTGATTCACGATGGAT 3598
 RESULT 9
 ID ACC72052 standard; DNA; 9695 BP.
 XX
 AC ACC72052;
 DT 08-JUL-2003 (first entry)
 XX
 DE BCU0205B gene #SEQ ID 81.
 XX
 KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
 KW drug discovery; clinical medicine; forensic medicine; gene;
 KW chromosome 5q33.3; ds.
 XX
 OS Homo sapiens.
 XX
 EN WO2003029421-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-US031287.
 XX
 PR 03-OCT-2001; 2001US-0326526P.
 PR 14-MAY-2002; 2002US-00144194.
 XX
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX
 PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
 XX
 DR WPI; 2003-381623/36.
 DR P-PSDB; ABR58318.
 XX
 PT New isolated human differentially-regulated breast cancer polynucleotide
 PT and polypeptide, useful for diagnosing, staging, prognosticating,
 PT preventing and/or treating diseases and conditions relating to breast
 PT cancer.


```
XX PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.
XX CC The invention relates to isolated polynucleotides which are
XX CC differentially-regulated in breast cancer. The methods and compositions
XX CC of the present invention are useful for detecting, diagnosing, staging,
XX CC monitoring, prognosticating, preventing and/or treating diseases and
XX CC conditions relating to breast cancer, and may be used in gene therapy or
XX CC antisense therapy. They can also be used in research, drug discovery,
XX CC clinical medicine and forensic medicine. Sequences given in records
XX CC ACC72012-ACC72074 represent polynucleotides of the invention that are
XX CC differentially-regulated in breast cancer. NOTE: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.01e-71 Length: 9695
Score: 613.00 Matches: 112
Percent Similarity: 89.40% Conservativity: 23
Best Local Similarity: 74.17% Mismatches: 16
Query Match: 78.39% Indels: 0
DB: 7 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x ACC72052 (1-9695)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3783 TACATCCGACGCACTTCTCCCTCTCGAATGTGACACGACATCTTGGAGTTACGAATAAA 3842
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3843 GAGTTTAAACATAGCAACACCCAGCACACAAAGTACTACTTGGCAGTGGACCCCGTGTC 3902
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3903 GGCTCGCTCTACGTGTCGACCCACCAACAGCAGGAGAAATCTACCGCGTCAAGTCTCTGAGT 3962
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3963 GGAACCAAGACCTGCTGGCAATTCGGAATGTGCGAGGACGGAGAGACAGTGTCTA 4022
QY 81 ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4023 CCCTTTGATGAAGCCCTCGGGGATGGAGGAGGAGCCATAGATGCAACCCCTGATGAGC 4082
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4083 CCAGAGGTATTGCAGTAGACAGAAATGGGCTATGTACTTGTGATGCCACCATGATC 4142
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4143 CGGAAGGTTGACCAAGATGAATCATCTCCACCTGCTGGCTCCATGACCTCACTGCC 4202
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4203 GTCCGGCCGCTGAGCTGTGATCCAGCATGGAT 4235

RESULT 10
AAS14089
ID AAS14089 standard; DNA; 9729 BP.
XX AC AAS14089;
XX CC
XX DT 18-DEC-2001 (first entry)
XX DE Human FCTR3f DNA sequence.
XX KW Human; FCTR3; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
XX KW astrocytoma; congenital neonatal alloimmune thrombocytopenia; infection;
XX KW neurological disorder; neurodegenerative disorders; nerve trauma;
```

```
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greenou type I; Corneal dystrophy-lattice type I;
KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW anti-allergic; antiasthmatic; antiinfertility; antiinflammatory;
KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3a; neurestin-like protein; FCTR3f.
XX OS Homo sapiens.
XX PH
XX FT 5'UTR Location/Qualifiers
XX FT 1..209
XX FT /*tag= b
XX FT CDS 210..8384
XX FT /*tag= a
XX FT /*product= "Human FCTR3f"
XX FT 3'UTR 8385..9729
XX FT /*tag= c
XX PN WO20016747-A2.
XX PD 13-SEP-2001.
XX PF 05-MAR-2001; 2001WO-US0007160.
XX PR 03-MAR-2000; 2000US-0186592P.
XX PR 03-MAR-2000; 2000US-0186718P.
XX PR 06-MAR-2000; 2000US-0187293P.
XX PR 06-MAR-2000; 2000US-0187294P.
XX PR 17-MAR-2000; 2000US-0190400P.
XX PR 07-APR-2000; 2000US-0196018P.
XX PR 03-JAN-2001; 2001US-0259548P.
XX PA (CURA-) CURAGEN CORP.
XX PI Vernet CM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
XX PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX DR WPI; 2001-596837/67.
XX DR P-PSDB; AAU08681.
XX PT Novel polypeptides designated as FCTR3 polypeptides, useful in detection,
XX PT prevention and treatment of a broad range of pathological states.
XX PS Claim 9; Page 37-39; 215pp; English.
XX CC The invention relates to human FCTR3 polypeptides, FCTR1-FCTR7, and the
XX CC nucleic acids encoding them. These sequences are useful for the treatment
XX CC or prevention of numerous disorders including myelogenous leukaemia,
XX CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
XX CC alloimmune thrombocytopenia, neurological disorders, neurodegenerative
XX CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
XX CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
XX CC myelodysplastic syndrome, mental health conditions, immunological
XX CC disorders, allergy and infection, bronchial asthma, Avellino type
XX CC eosinophilia, lung diseases, reproductive disorders, infertility, male
XX CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
XX CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
XX CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
XX CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
XX CC Corneal dystrophy-Greenou type I, Corneal dystrophy-lattice type I and
XX CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
XX CC the FCTR3a homologue FCTR3f
XX SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;
```

Alignment Scores:
 Pred. No.: 1,02e-71 Length: 9729
 Score: 613.00 Matches: 112
 Percent Similarity: 89.40% Conservative: 23
 Best Local Similarity: 74.17% Mismatches: 16
 Query Match: 78.39% Indels: 0
 DB: 5 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AAS14089 (1-9729)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
 DB 3831 TACATCCGACGATCTTCCCTTCGAAATGTGACACGATCTTGGAGTTACGAATAAA 3890

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
 DB 3891 GAGTTTAAACATAGCAACACACACGACACACAGTACTACTTGGCAGTGGACCCCGTGTCC 3950

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrVal 60
 DB 3951 GGCTCGCTCTAGTGTCCGACACACACAGAGAGATCTACCGGTCAAGTCTCTGACT 4010

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
 DB 4011 GGAACCAAGACCTGGCTGGGAATTCGAAATGTGCGCAGGACGGGAGCAGTGTCTA 4070

QY 81 PropheAspThrArgCysGlyAspGlyGlyLeuAlaThrGluAlaThrLeuThrAsn 100
 DB 4071 CCTTTGATGAAGCCCGCTGCGGGGATGGAGGAGCCATAGATCAACCCCTGATGAGC 4130

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
 DB 4131 CCGAGAGGTATTGCCAGTACAGAAATGGCTCATGTACTTTGTCGATGCCACCATGATC 4190

QY 121 ArgArgIleAspGlnAsnGlyIleLeuSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 DB 4191 CGAAGGTTGACCAAGATGGAATCATCTCACCTGTGGCTTCCCAATGACCTCACTGCC 4250

QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
 DB 4251 GTCCGCGCGCTGAGTGTGATTCCAGCATGGAT 4283

RESULT 11
 ADB32028
 ID ADB32028 standard; cDNA; 9729 BP.
 XX AC ADB32028;
 XX DT
 XX DE
 XX DE Human FCTR3f cDNA.
 XX KW Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
 KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
 KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
 KW renal cell carcinoma; melanoma; clear cell carcinoma;
 KW granular cell carcinoma; neurological disorder;
 KW neurodegenerative disorder; nerve trauma;
 KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
 KW Gardner syndrome; mental health condition; immunological disorder;
 KW allergy; asthma; lung disease; reproductive disorder; deafness;
 KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
 KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
 KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
 KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
 KW Lattice corneal dystrophy.
 XX OS Homo sapiens.
 XX PN US2003087816-A1.
 XX PD 08-MAY-2003.
 XX XX

PF 05-MAR-2001; 2001US-00800198.
 PR 03-MAR-2000; 2000US-0186592P.
 XX (VERM/) VERMET C.
 PA (FERN/) FERNANDES E.
 PA (SHIM/) SHIMKETS R.
 PA (HERR/) HERRMANN J.
 PA (MAJU/) MAJUMDER K.
 PA (MACD/) MACDOUGALL J.
 PA (MISH/) MISHRA V.
 PA (MEZE/) MEZES P S.
 PA (RAST/) RASTELLI L.
 XX Vernet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
 PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
 DR P-PSDB; ADB32029.
 XX WPI; 2003-625633/59.
 DR P-PSDB; ADB32029.
 XX New FCTR3f polypeptide and encoding polynucleotide, useful for preventing
 PT or treating FCTR3f-related disorders, such as cancer, autoimmune,
 PT neurodegenerative, gastrointestinal, reproductive and inflammatory
 PT diseases.
 XX Claim 9; Page 34-37; 155pp; English.
 XX The invention relates to FCTR3f polypeptides and the polynucleotides
 CC encoding them. The sequences of the invention are useful for the
 CC manufacture of a medicament for diagnosing and treating disorders
 CC associated with the FCTR3f polypeptide, such as colorectal cancer,
 CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
 CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
 CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
 CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
 CC cell and granular cell carcinomas, neurological disorders,
 CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
 CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
 CC conditions, immunological disorders, allergy and infection, asthma, lung
 CC diseases, male and female reproductive disorders, deafness, glycoprotein
 CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
 CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
 CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
 CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
 CC represents cDNA encoding an FCTR3f polypeptide of the invention.
 XX Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 1,02e-71 Length: 9729
 Score: 613.00 Matches: 112
 Percent Similarity: 89.40% Conservative: 23
 Best Local Similarity: 74.17% Mismatches: 16
 Query Match: 78.39% Indels: 0
 DB: 5 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x ADB32028 (1-9729)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
 DB 3831 TACATCCGACGATCTTCCCTTCGAAATGTGACACGATCTTGGAGTTACGAATAAA 3890

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
 DB 3891 GAGTTTAAACATAGCAACACACACGACACACAGTACTACTTGGCAGTGGACCCCGTGTCC 3950

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgValPheLysIleLysSerThrVal 60
 DB 3951 GGCTCGCTCTAGTGTCCGACACACACAGAGAGATCTACCGGTCAAGTCTCTGACT 4010

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
 DB 4011 GGAACCAAGACCTGGCTGGGAATTCGAAATGTGCGCAGGACGGGAGCAGTGTCTA 4070

QY 81 ProPheAspThrArgCysGlyAspGlyValAlaThrGluAlaThrLeuThrAsn 100
 Db 4071 CCCTTTGATGAAGCCCGCTGGGGATGGAGGAAGCCATAGATCAACCTGATGAGC 4130
 QY 101 ProArgGlyLeuThrValAspLysPheGlyLeuLeuTyrPheValAspGlyThrMetIle 120
 Db 4131 CCGAGAGTATTGAGTAGACAGAAATGGGCTCAATGACTTTCGATGCCACCATGATC 4190
 QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 Db 4191 CGGAAGGTTCACCAAGATGAATCATCTCCACCTGCTGGCTCCATGACCTCACTGCC 4250
 QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
 Db 4251 GTCCGGCCGCTGAGCTGTGATTCAGCATGGAT 4283

RESULT 12
 AAS14085
 ID AAS14085 standard; DNA; 9826 BP.
 AC AAS14085;
 XX 18-DEC-2001 (first entry)
 XX Human FCTR3b DNA sequence.

KW Human; FCTRX; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
 KW astrocytoma; congenital neonatal alloimmune thrombocytopenia; infection;
 KW neurological disorder; neurodegenerative disorders; nerve trauma;
 KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
 KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
 KW mental health condition; immunological disorder; allergy; infertility;
 KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
 KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
 KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
 KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
 KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
 KW Corneal dystrophy-Greenow type I; Corneal dystrophy-lattice type I;
 KW Reis-Bucklers corneal dystrophy; cystostatic; immunosuppressive;
 KW anti-allergic; antiasthmatic; antiinfertility; antiinflammatory;
 KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
 KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
 KW gene therapy; FCTR3b; neurestin-like protein.

XX Homo sapiens.

XX Key
 FH 5'UTR 1..279 Location/Qualifiers
 FT /*tag= b
 FT 280..8481
 FT /*tag= a
 FT /product= "Human FCTR3b"
 FT 8482..9826
 FT /*tag= c

XX WO200166747-A2.
 XX 13-SEP-2001.
 XX 05-MAR-2001; 2001WO-US007160.
 XX 03-MAR-2000; 2000US-0186592P.
 XX 03-MAR-2000; 2000US-0186718P.
 XX 06-MAR-2000; 2000US-0187293P.
 XX 17-MAR-2000; 2000US-0187294P.
 XX 07-APR-2000; 2000US-0190400P.
 XX 03-JAN-2001; 2001US-0259548P.
 XX (CURA-) CURAGEN CORP.
 XX Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;

PI MacDougall J, Mishra V, Mezes PS, Rastelli L;
 XX WPI; 2001-596837/67.
 DR P-PSDB; AAU08680.
 XX Novel polypeptides designated as FCTRX polypeptides, useful in detection,
 PT prevention and treatment of a broad range of pathological states.
 XX Claim 9; Page 33-35; 215pp; English.

XX The invention relates to human FCTRX polypeptides, FCTR1-FCTR7, and the
 CC nucleic acids encoding them. These sequences are useful for the treatment
 CC or prevention of numerous disorders including myelogenous leukaemia,
 CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
 CC alloimmune thrombocytopenia, neurological disorders, neurodegenerative
 CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
 CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
 CC myelodysplastic syndrome, mental health conditions, immunological
 CC disorders, allergy and infection, bronchial asthma, Avellino type
 CC eosinophilia, lung diseases, reproductive disorders, infertility, male
 CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
 CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
 CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
 CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
 CC Corneal dystrophy-Greenow type I, Corneal dystrophy-lattice type I and
 CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
 CC FCTR3b, a neurestin-like protein

XX SQ Sequence 9826 BP; 2567 A; 2643 C; 2479 G; 2137 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.03e-71 Length: 9826
 Score: 613.00 Matches: 112
 Percent Similarity: 89.40% Conservations: 23
 Best Local Similarity: 74.17% Mismatches: 16
 Query Match: 78.39% Indels: 0
 DB: 5 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AAS14085 (1-9826)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnIlys 20
 Db 3928 TACATCCGAGCATCTTCCCTCTCGAAATGTGCAGCATCTTGGAGTTACGAAATAA 3987

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
 Db 3988 GAGTTTAAACATAGCAACAACCCAGCACACAAGTACTACTTGGCAGTGGACCCCGTCTCC 4047

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
 Db 4048 GCTCGCTCTACGTGTCGACACCAACAGCAGGAGAAATCTACCCGCTCAAGTCTCTGAGT 4107

QY 61 ValValIlysAspLeuValIlysAsnSerGluValAlaGlyThrGlyAspGlnCysLeu 80
 Db 4108 GGAACCAAGACCTGGCTGGGAATTCGAAAGTTGTGCGAGGACGGGAGACATGTCTA 4167

QY 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
 Db 4168 CCCTTTGATGAAGCCCGCTGGGGATGGAGGAAGGCCATAGATCAACCTGATGAGC 4227

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuLeuTyrPheValAspGlyThrMetIle 120
 Db 4228 CCGAGAGTATTGAGTAGACAGAAATGGGCTCATGACTTTCGATGCCACCATGATC 4287

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
 Db 4288 CGGAAGGTTCACCAAGATGAATCATCTCCACCTGCTGGCTCCATGACCTCACTGCC 4347

QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
 Db 4348 GTCCGGCCGCTGAGCTGTGATTCAGCATGGAT 4380

RESULT 13

```
ADB32023
ID ADB32023 standard; cDNA; 9826 BP.
XX
AC ADB32023;
XX
XX 04-DEC-2003 (first entry)
XX
DE Human FCTR3b cDNA.
XX
XX Human; gene; ss; FCTR3b; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW
XX
XX Homo sapiens.
XX
XX US2003087816-A1.
XX
XX 08-MAY-2003.
XX
XX 05-MAR-2001; 2001US-00800198.
XX
XX 03-MAR-2000; 2000US-0186592P.
XX
XX (VERM/) VERMET C.
XX (FERN/) FERNANDES E.
XX (SHIM/) SHIMKETS R.
XX (HERR/) HERRMANN J.
XX (MAJU/) MAJUMDER K.
XX (MACD/) MACDOUGALL J.
XX (MISH/) MISHRA V.
XX (MEZE/) MEZES P S.
XX (RAST/) RASTELLI L.
XX
XX Vernet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
PI MacDougall J, Mishra V, Mezes PS, Rastelli L;
XX
XX WPI; 2003-625633/59.
XX P-PSDB; ADB32024.
XX
XX New FCTR3 polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTR3-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.
XX
XX Claim 9; Page 29-32; 155pp; English.
XX
XX The invention relates to FCTR3 polypeptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTR3 polypeptide, such as colorectal cancer,
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, neurological disorders,
CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
CC conditions, immunological disorders, allergy and infection, asthma, lung
CC diseases, male and female reproductive disorders, deafness, glycoprotein
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC
```

```
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTR3 polypeptide of the invention.
XX
SQ Sequence 9826 BP; 2567 A; 2645 C; 2476 G; 2138 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.03e-71 Length: 9826
Score: 613.00 Matches: 112
Percent Similarity: 89.40% Conservative: 23
Best Local Similarity: 74.17% Mismatches: 16
Query Match: 78.39% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x ADB32023 (1-9826)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleuGluLeuArgAsnLys 20
DB 3928 TATATCCGACGCATCTTCCCTCTCGAATGTGACCAAGCATCTTGAGTTAGAAATAA 3987
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPProMetSer 40
DB 3988 GAGTTTAAACATAGCAACACCAACACACACAGTACTACTTGGCAGTGGACCCCGTGTC 4047
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
DB 4048 GGCCTGCTCTAGTGTCCGACACCAACACAGCAGAGAATCTACCGCTCAAGTCTCTGAGT 4107
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
DB 4108 GGAACCAAGACCTGGCTGGGAATTCGGAAGTTGTGCGAGGACGGGAGCAGTGTCTA 4167
QY 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
DB 4168 CCTTTGATGAAGCCCGCTGCGGGATGGAGGAGCCATAGATCAACCTGTATGAGC 4227
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
DB 4228 CCGAGAGGTATTGCAGTAGACAAGAATGGGCTCATGTACTTTGTCTGATGCCACATGATC 4287
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
DB 4288 CGAAGGTTGACCAAGTGAATCATCTCCACCTGCTGGCTCCCAATGACCTCACTGTC 4347
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
DB 4348 GTCCGGCCCGCTGAGCTGTGATTCACCATGGAT 4380

RESULT 14
AAL60066
ID AAL60066 standard; cDNA; 4245 BP.
XX
AC AAL60066;
XX
XX 27-AUG-2003 (first entry)
XX
XX Human Pc099 cDNA.
XX
XX Human; differentially regulated protein; prevention; therapy; vaccine;
KW prostate cancer; gene therapy; Pc099; teneurin-2; chromosome 5q34; gene;
XX ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..4056
XX /tag= a
XX /product= "Human Pc099 protein"
XX /note= "No start codon"
XX /partial
XX
XX WO2003040331-A2.
XX
XX 15-MAY-2003.
XX
```

```
XX PF 07-NOV-2002; 2002WO-US035563.
XX PF
XX PR 07-NOV-2001; 2001US-0331041P.
XX PR 07-NOV-2001; 2001US-0331042P.
XX PR 18-DEC-2001; 2001US-0340251P.
XX PR 07-JAN-2002; 2002US-0344791P.
XX PR
XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX PI
XX PI Sun Z, Li X, Jay G, Kovacs KF, Fan W;
XX DR WPI; 2003-449451/42.
XX DR P-PSDB; AAO29571.
XX XX
XX PT New polynucleotide for diagnosing, staging, monitoring, prognosticating,
XX PT preventing or treating, or determining the predisposition to, diseases or
XX PT conditions such as prostate cancer, and for research or forensic science.
XX XX
XX PS Claim 29; Page 158-163; 100pp; English.
XX XX
XX CC The present invention relates to novel differentially regulated genes and
XX CC polypeptides encoded by them. Sequences of the invention are useful in
XX CC diagnosing, staging, monitoring, prognosticating, preventing, treating or
XX CC determining the predisposition to diseases or conditions such as prostate
XX CC cancer. They may be used as molecular markers, drug targets, vaccines, in
XX CC gene therapy, research, clinical medicine or forensic science. The
XX CC present sequence is a differentially regulated prostate cDNA, Pc099 which
XX CC codes for teneurin-2. Pc099 gene is located on chromosome 5q34
XX XX
XX SQ Sequence 4245 BP; 1035 A; 1132 C; 1115 G; 963 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.36e-66 Length: 4245
Score: 571.50 Matches: 108
Percent Similarity: 84.77% Conservative: 20
Best Local Similarity: 71.52% Mismatches: 16
Query Match: 73.08% Indels: 7
DB: 7 Gaps: 1

US-10-029-020-14_COPY_1250_1400 (1-151) x AAL60066 (1-4245)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 2365 TACATCGACGATCTTCCCTCTCGAAATGTGACGAGATCTTGGAGTTACGCAAC--- 2421
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 2422 -----AACCCAGCACACAAAGTACTTACTTGGCAGTGGACCCCGTGCTCC 2463
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 2464 GGTCGCTCTACGTGTCCGACCAACACAGCAGGAGAAATCTACCGCGTCAAGTCTCTGAGT 2523
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 2524 GGAACCAAGACCTGGTGGATTCGGAGTTGTGCGAGGAGCGGAGAGCAGTGTCTA 2583
QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 2584 CCCTTTGATGAAGCCCTCGGGGGATGGAGGGAAGGCCATAGATCAACCCCTGATGAGC 2643
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 2644 CCAGAGAGTATTGAGTAGACAAAGATGGGTCATGTACTTGTGATGCCCATGATC 2703
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 2704 CGGAGGTTGACCAAGATGAATCATCTCCACCTGCTGGGCTCCATGACCTCACTGCC 2763
QY 141 AlArgProLeuSerCysAspSerValMetAsp 151
Db 2764 GTCGGCGCGTGAGCTGTGATTCAGCATGGAT 2796
```

```
RESULT 15
ABK92230
ID ABK92230 standard; DNA; 12879 BP.
XX AC
XX AC ABK92230;
XX DT
XX DT 15-AUG-2002 (first entry)
XX XX
XX XX Prostate cancer-associated DNA sequence #116.
XX XX
XX XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
XX KW gene therapy; gene; ds.
XX KM Mammalia.
XX OS
XX XX
XX PN WO200230268-A2.
XX XX
XX PD 18-APR-2002.
XX XX
XX PF 12-OCT-2001; 2001WO-US032045.
XX XX
XX PR 13-OCT-2000; 2000US-00687576.
XX PR 08-DEC-2000; 2000US-00733288.
XX PR 08-DEC-2000; 2000US-00733742.
XX PR 24-JAN-2001; 2001US-0263957P.
XX PR 16-MAR-2001; 2001US-0276791P.
XX PR 16-MAR-2001; 2001US-0276888P.
XX PR 06-APR-2001; 2001US-0281922P.
XX PR 24-APR-2001; 2001US-0286214P.
XX PR 30-APR-2001; 2001US-00847046.
XX PR 04-MAY-2001; 2001US-0288589P.
XX XX
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX XX
XX PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX XX
XX DR WPI; 2002-471335/50.
XX DR P-PSDB; ABG61913.
XX XX
XX PT Detecting a prostate cancer-associated transcript in a cell in a patient,
XX PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
XX PT by determining if prostate cancer-associated genes are expressed in a
XX PT prostate tissue.
XX XX
XX PS Claim 22; Page 394-397; 436pp; English.
XX XX
XX CC The present invention relates to methods of detecting a prostate cancer-
XX CC associated transcript in a cell from a patient. The method comprises
XX CC contacting a biological sample from the patient with prostate cancer-
XX CC associated polynucleotides (designated PC genes) that selectively
XX CC hybridise to a sequence that is at least 80% identical to them. The
XX CC prostate cancer-associated polynucleotide sequences are differentially
XX CC expressed in prostate tumour tissue or in prostate cancer and are derived
XX CC from the tissues of various organisms such as humans or other mammals
XX CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
XX CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
XX CC associated genes are useful for diagnosing or treating prostate cancer,
XX CC as well as for identifying modulators of prostate cancer or agents that
XX CC inhibit prostate cancer. The nucleic acid sequences are particularly
XX CC useful in gene therapy, as a vaccine or in antisense applications.
XX CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
XX CC sequences
XX XX
XX SQ Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.46e-61 Length: 12879
Score: 534.50 Matches: 100
Percent Similarity: 82.78% Conservative: 25
Best Local Similarity: 66.23% Mismatches: 19
Query Match: 68.35% Indels: 7
DB: 6 Gaps: 1
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 202.425 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_1100_1200

Perfect score: 549

Sequence: 1 RUPRKFKAAPDLISYFIWD.....IQSGILHKGNGENQFVSQQP 101

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-Q=/cgn2_1/USPTO.spool/US10029020/runat_06082004/112218_29331/app_query.fasta_1.3519
-DB=Published Applications NA -OFMT=Fastap -SURFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020@cgn_1_1 2156 @runat_06082004/112218_29331
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
|------------|-------|-------------|--------|----|-------------|

ALIGNMENTS

RESULT 1

US-10-383-201-43

; Sequence 43, Application US/10383201

; Publication No. US20040029236A1

; GENERAL INFORMATION:

; APPLICANT: Alsbrook II, John et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES,

; FILE REFERENCE: 21402-568A

; CURRENT APPLICATION NUMBER: US/10/383,201

; CURRENT FILING DATE: 2003-03-06

; PRIOR APPLICATION NUMBER: 10/029020

; PRIOR FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/365,984

; PRIOR FILING DATE: 2002-03-20

; PRIOR APPLICATION NUMBER: 60/372,022

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/389,143

; PRIOR FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 60/391,779

; PRIOR FILING DATE: 2002-06-26

; PRIOR APPLICATION NUMBER: 60/410,755

; PRIOR FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: 60/412,957

| | | | | | | |
|----|-------|-------|---------|----|---------------------|--------------------|
| 1 | 549 | 100.0 | 8354 | 13 | US-10-383-201-43 | Sequence 43, Appl |
| 2 | 549 | 100.0 | 8354 | 13 | US-10-029-020-13 | Sequence 13, Appl |
| 3 | 538.5 | 98.1 | 8355 | 13 | US-10-383-201-55 | Sequence 55, Appl |
| 4 | 538.5 | 98.1 | 8438 | 13 | US-10-042-865-1 | Sequence 1, Appl |
| 5 | 453 | 82.5 | 8473 | 17 | US-10-038-854-39 | Sequence 39, Appl |
| 6 | 453 | 82.5 | 8487 | 17 | US-10-038-854-41 | Sequence 41, Appl |
| 7 | 453 | 82.5 | 8645 | 17 | US-10-038-854-37 | Sequence 37, Appl |
| 8 | 453 | 82.5 | 8675 | 17 | US-10-038-854-35 | Sequence 35, Appl |
| 9 | 426 | 77.6 | 8409 | 9 | US-09-808-602-79 | Sequence 79, Appl |
| 10 | 426 | 77.6 | 8409 | 10 | US-09-800-198-67 | Sequence 67, Appl |
| 11 | 417 | 76.0 | 8797 | 9 | US-09-808-602-74 | Sequence 74, Appl |
| 12 | 417 | 76.0 | 8797 | 9 | US-09-808-602-77 | Sequence 77, Appl |
| 13 | 417 | 76.0 | 8797 | 10 | US-09-800-198-62 | Sequence 62, Appl |
| 14 | 417 | 76.0 | 8797 | 10 | US-09-800-198-65 | Sequence 65, Appl |
| 15 | 416 | 75.8 | 6560 | 9 | US-09-808-602-76 | Sequence 76, Appl |
| 16 | 416 | 75.8 | 6560 | 10 | US-09-800-198-64 | Sequence 64, Appl |
| 17 | 416 | 75.8 | 9058 | 16 | US-10-144-194A-79 | Sequence 79, Appl |
| 18 | 416 | 75.8 | 9695 | 16 | US-10-144-194A-81 | Sequence 81, Appl |
| 19 | 411 | 74.9 | 12880 | 16 | US-10-295-027-927 | Sequence 927, Appl |
| 20 | 407 | 74.1 | 9729 | 9 | US-09-808-602-12 | Sequence 12, Appl |
| 21 | 407 | 74.1 | 9729 | 10 | US-09-800-198-12 | Sequence 12, Appl |
| 22 | 407 | 74.1 | 9826 | 9 | US-09-808-602-7 | Sequence 7, Appl |
| 23 | 407 | 74.1 | 9826 | 10 | US-09-800-198-7 | Sequence 7, Appl |
| 24 | 406 | 74.0 | 8689 | 9 | US-09-808-602-78 | Sequence 78, Appl |
| 25 | 406 | 74.0 | 8689 | 10 | US-09-800-198-66 | Sequence 66, Appl |
| 26 | 341 | 62.1 | 466 | 16 | US-10-125-968-61 | Sequence 143, Appl |
| 27 | 323 | 58.8 | 8575 | 13 | US-10-072-012-143 | Sequence 143, Appl |
| 28 | 221 | 40.3 | 534 | 13 | US-10-027-632-49758 | Sequence 49758, A |
| 29 | 221 | 40.3 | 534 | 13 | US-10-027-632-49759 | Sequence 49759, A |
| 30 | 221 | 40.3 | 534 | 16 | US-10-027-632-49758 | Sequence 49758, A |
| 31 | 221 | 40.3 | 534 | 16 | US-10-027-632-49759 | Sequence 49759, A |
| 32 | 125 | 22.8 | 310 | 15 | US-10-029-386-16277 | Sequence 16277, A |
| 33 | 125 | 22.8 | 559 | 15 | US-10-029-386-2577 | Sequence 1867, Ap |
| 34 | 75 | 13.7 | 71132 | 13 | US-10-087-192-1867 | Sequence 2, Appl |
| 35 | 73.5 | 13.4 | 3673778 | 15 | US-10-312-841-2 | Sequence 9600, Ap |
| 36 | 73 | 13.3 | 1908 | 16 | US-10-040-862-9600 | Sequence 9600, Ap |
| 37 | 73 | 13.3 | 1908 | 16 | US-10-057-4758-9600 | Sequence 9600, Ap |
| 38 | 73 | 13.3 | 1908 | 16 | US-10-154-8848-9600 | Sequence 6, Appl |
| 39 | 73 | 13.3 | 2516 | 13 | US-09-795-651-6 | Sequence 4, Appl |
| 40 | 73 | 13.3 | 3023 | 9 | US-09-900-236-4 | Sequence 4, Appl |
| 41 | 73 | 13.3 | 3023 | 15 | US-10-066-443-4 | Sequence 4, Appl |
| 42 | 73 | 13.3 | 3023 | 15 | US-10-256-889-4 | Sequence 4, Appl |
| 43 | 73 | 13.3 | 3023 | 16 | US-10-439-799-4 | Sequence 4, Appl |
| 44 | 73 | 13.3 | 6303 | 13 | US-10-342-887-1424 | Sequence 1424, Ap |
| 45 | 73 | 13.3 | 6303 | 13 | US-10-172-118-1424 | Sequence 1424, Ap |


```
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:
Pred. No.: 5,71e-69 Length: 8355
Score: 538.50 Matches: 101
Percent Similarity: 99.02% Conservative: 0
Best Local Similarity: 99.02% Mismatches: 0
Query Match: 98.09% Indels: 1
DB: 13 Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-383-201-55 (1-8355)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3301 CGCCTCTTCAGGAAGTGTTCGCTGCAGCCCGCCAGACCTGCTATTATTATTTCAITTTGGGAC 3360
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3361 AAGACAGACGCTACACACAGAGAGGTGTTTGGGCTTTTCAGAACCTTTGTTCCGTGGGT 3420
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 3421 TATGAATATGAATCTTCGCCAGATCTAATCTGTGGGAAAAGAACACAGTGTCTGCAG 3480
QY 61 GlyTyrGluLeuAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db 3481 GGCTATGAATTCATGCGTCCAGCTTGGAGGATGGAGCCCTAGACAAACATCATGCCCTC 3540
QY 81 AsnIleGlnSer---GlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGln 99
Db 3541 AACATTCAAGTGGTGGCATCTGCACAAAGGGAATGGGAGAACCCAGTTTGTGTCTCAG 3600
QY 100 GlnPro 101
Db 3601 CAGCCT 3606

RESULT 4
US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tcherenev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
```

```
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1

Alignment Scores:
Pred. No.: 5,79e-69 Length: 8438
Score: 538.50 Matches: 101
Percent Similarity: 99.02% Conservative: 0
Best Local Similarity: 99.02% Mismatches: 0
Query Match: 98.09% Indels: 1
DB: 13 Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-042-865-1 (1-8438)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3311 CGCCTCTTCAGGAAGTGTTCGCTGCAGCCCGCCAGACCTGCTCTATTATTATTTCAITTTGGGAC 3370
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3371 AAGACAGACGCTACACACAGAGGTGTTTGGGCTTTTCAGAACCTTTGTTCCGTGGGT 3430
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 3431 TATGAATATGAATCTTCGCCAGATCTAATCTGTGGGAAAAGAACACAGTGTCTGCAG 3490
QY 61 GlyTyrGluLeuAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db 3491 GGCTATGAATTCAGCGCTCCAGCTTGGAGGATGGAGCCCTAGACAAACATCATGCCCTC 3550
QY 81 AsnIleGlnSer---GlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGln 99
Db 3551 AACATTCAAGTGGTGGCATCTGCACAAAGGGAATGGGAGAACCCAGTTTGTGTCTCAG 3610
QY 100 GlnPro 101
Db 3611 CAGCCT 3616

RESULT 5
US-10-038-854-39
; Sequence 39, Application US/10038854
; Publication No. US20040022781A1
```

Qy 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20

```
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-41

Alignment Scores:
Pred. No.: 3,67e-56 Length: 8487
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-038-854-41 (1-8487)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3146 AGACTCTTCCAAAAGTGGTTCTCCTGCCTCACCACAACTTGGCCTATATCTTTTCATATGGGAT 3205
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3206 AAAACAGATGATATATATCAGAAAGTCTATGCTATCTGAGCTGTGTGTCAGTTGGA 3265
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGln 60
Db 3266 TATGAGTATGAGTCGTGTTGGACCTGACTCTCTGGGAAAAGAGGACTGCCATTCTGCAG 3325
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 3326 GGCTATGATGATGATGCTGCTCAACATGGGTGGCTGGACATTAGATAAACATCAGTGTG 3385
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3386 GATGTACAGAACGGTATATCTACTGACAGGAAACGGGGAAACCCAGTTCACTCTCCAGCAG 3445
QY 101 Pro 101
Db 3446 CCT 3448

RESULT 7
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytex, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malvankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gorman, Steven K
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
```

```
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 8645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-37

Alignment Scores:
Pred. No.: 3,77e-56 Length: 8645
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-038-854-37 (1-8645)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3322 AGACTCTTCCAAAAGTGGTTCTCCTGCCTCACCACAACTTGGCCTATATCTTTTCATATGGGAT 3381
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3382 AAAACAGATGATATATATCAGAAAGTCTATGCTATCTGAGCTGTGTGTCAGTTGGA 3441
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGln 60
Db 3442 TATGAGTATGAGTCGTGTTGGACCTGACTCTCTGGGAAAAGAGGACTGCCATTCTGCAG 3501
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 3502 GGCTATGATGATGATGCTGCTCAACATGGGTGGCTGGACATTAGATAAACATCAGTGTG 3561
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3562 GATGTACAGAACGGTATATCTACTGACAGGAAACGGGGAAACCCAGTTCACTCTCCAGCAG 3621
QY 101 Pro 101
Db 3622 CCT 3624
```

```

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-038-854-35 (1-8675)

Qy      1 ArgLeuPheArgLysTyrPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp    20
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     3334 AGACTCTCCAAAGAGTGTTCTCGCCTCACCAACCTGGCTATACCTTCATATGGGAT   3393

Qy     21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly    40
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     3394 AAAACAGATGCATAAATCAGAAAGTCTATGGTCTATCTGAAGCTGTGTGTCAATTGGA   3453

Qy     41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln    60
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     3454 TATGAGTATGAGTCGTGTGGACCTGACTCTGTGGGAAAAGAGACTGCCCATTCGCGAG   3513

Qy     61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu    80
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     3514 GCGCTATCAATTGGATCGTCGTCACAATGGTGGCTGGACATTAGATAAACATCACGTGCTG   3573

Qy     81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln   100
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     3574 GATGTACAGAACGGGTACTGTACAGGGAAACGGGGAAAAACCAGTTCATCTCCAGCAG   3633

Qy      101 Pro 101
       |||
Db     3634 CCT 3636

RESULT 9
US-09-808-602-79
; Sequence 79, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herрман, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: Macdougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15965-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-808-602-79

Alignment Scores:
Pred. No.:          3 99e-52           Length:         8409
Score:             426.00            Matches:         74
Percent Similarity: 89.00%           Conservatives:    15
Best Local Similarity: 74.00%        Mismatches:      11
Query Match:       77.60%            Indels:          0
DB:                9                 Gaps:            0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-808-602-79 (1-8409)

```

| | | | |
|----|------|--|------|
| QY | 2 | LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleIleTIPAspLys | 21 |
| | |:::..... | |
| Db | 3409 | CTATTTCAAAAATCATTTCTGGCATCTCCCAATCTGGCTTATCATTCATCGGACAAA | 3468 |
| QY | 22 | ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValClyTyr | 41 |
| | |:::..... | |
| Db | 3469 | ACAGATGCATATCGTCAGAAAGTTTATGGCTGTGCAGATCTGAGTTTCTGTGGGTATT | 3528 |

```
QY 42 GluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGlnGly 61
Db 3529 GAATATGAGACTTGTCCAGCTTTGATTTCTGTGGAGAAAAGAGACTGGCTGTGCAAGGA 3588
QY 62 TyrGluIleAspAlaSerLysLeuGlyTyrSerLeuAspLysHisAlaLeuAsn 81
Db 3589 TTTGAGCTAGATCCTTCCCAATCTAGGAGGATGCTTTGGATAAATCATCATGTACTGAAT 3648
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
Db 3649 GTCAAGAGTGTATATTGCACAAAGCAATGGAGAAATCAGTTTCTAACTCAGCAGCCA 3708

RESULT 10
US-09-800-198-67
; Sequence 67, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-800-198-67
```

```
Alignment Scores:
Pred. No.: 3,99e-52 Length: 8409
Score: 426.00 Matches: 74
Percent Similarity: 89.00% Conservative: 15
Best Local Similarity: 74.00% Mismatches: 11
Query Match: 77.60% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-800-198-67 (1-8409)

QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
Db 3409 CTATTTCAAATAATCATTTCTGGCATCTCCCAACTTGGCTTATACATTCATCTGGGACAAA 3468
QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
Db 3469 ACAGATGCATATGTCAGAGAGTTTATGGGTTGTGCAGATGCTGTAGTTTCTGTGGGTTT 3528
QY 42 GluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGlnGly 61
Db 3529 GAATATGAGACTTGTCCAGCTTTGATTTCTGTGGAGAAAAGAGACTGGCTGTGCAAGGA 3588
QY 62 TyrGluIleAspAlaSerLysLeuGlyTyrSerLeuAspLysHisAlaLeuAsn 81
Db 3589 TTTGAGCTAGATCCTTCCCAATCTAGGAGGATGCTTTGGATAAATCATCATGTACTGAAT 3648
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
Db 3649 GTCAAGAGTGTATATTGCACAAAGCAATGGAGAAATCAGTTTCTAACTCAGCAGCCA 3708

RESULT 11
US-09-800-602-74
; Sequence 74, Application US/09808602
```

```
Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-808-602-74

Alignment Scores:
Pred. No.: 9.47e-51 Length: 8797
Score: 417.00 Matches: 72
Percent Similarity: 88.00% Conservative: 16
Best Local Similarity: 72.00% Mismatches: 12
Query Match: 75.96% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-808-602-74 (1-8797)

QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
Db 3407 CTCTTCAGAGAGTCATTCAGGCTTCTCCCAACTAGCCCTACATTCATCTGGGACAG 3466
QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
Db 3467 ACAGATGCTTATGGCCAAAGGTTTATGGCTATCGGATGCTGTGTGTCTGTGGGTTT 3526
QY 42 GluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGlnGly 61
Db 3527 GAATATGAGACTTGTCCAGCTTCTATCTCTGTGGAGAAAAGAGAGCCCTGTCTTCAGGA 3586
QY 62 TyrGluIleAspAlaSerLysLeuGlyTyrSerLeuAspLysHisAlaLeuAsn 81
Db 3587 TTTGAGCTGAGCCCTTCCCAACTTGGAGGCTGTCTCTGGACAAACACACACCTCAAT 3646
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
Db 3647 GTCAAGAGCGGATACTATACAAAGGAGGAGGAGAGAACAGTTCTCTGACCCAGCAGCT 3706

RESULT 12
US-09-808-602-77
; Sequence 77, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
```

```
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 77
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-808-602-77

Alignment Scores:
Pred. No.: 9,47e-51 Length: 8797
Score: 417.00 Matches: 72
Percent Similarity: 88.00% Conservative: 16
Best Local Similarity: 72.00% Mismatches: 12
Query Match: 75.96% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-808-602-77 (1-8797)

QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyPheIleTrpAspLys 21
DB 3407 CTCITCCAGAGTCATTCAGGCTTCTCCACCTAGCTACACATTCATCTGGGACAAG 3466

QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
DB 3467 ACAGATGCTTATGGCCAAAGGGTTTATGGCTATCGATGCTGTGTGTCTGTGGGTTT 3526

QY 42 GluTyrGluSerCysProAspLeuLeuTrpGluLysArgThrValLeuGlnGly 61
DB 3527 GAAATATGAGACTGCCCGAGTCTCATCTGTGGGAGAAAGACAGCCCTGCTTCAGGGA 3586

QY 62 TyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeuAsn 81
DB 3587 TTCGAGCTGGACCTTCCAACTTGGAGGTGCTCCCTGGACAAACACACCCCTCAAT 3646

QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
DB 3647 GTGAAAGCGGAATACTACAAAGGACAGGGGAGAACCACTTCTGACCCAGCAGCCT 3706

RESULT 14
US-09-800-198-65
; Sequence 65, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 65
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-65

Alignment Scores:
Pred. No.: 9,47e-51 Length: 8797
Score: 417.00 Matches: 72
Percent Similarity: 88.00% Conservative: 16
Best Local Similarity: 72.00% Mismatches: 12
Query Match: 75.96% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-800-198-65 (1-8797)

QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyPheIleTrpAspLys 21
DB 3407 CTCITCCAGAGTCATTCAGGCTTCTCCACCTAGCTACACATTCATCTGGGACAAG 3466

QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
DB 3467 ACAGATGCTTATGGCCAAAGGGTTTATGGCTATCGATGCTGTGTGTCTGTGGGTTT 3526

QY 42 GluTyrGluSerCysProAspLeuLeuTrpGluLysArgThrValLeuGlnGly 61
DB 3527 GAAATATGAGACTGCCCGAGTCTCATCTGTGGGAGAAAGACAGCCCTGCTTCAGGGA 3586

QY 62 TyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeuAsn 81
DB 3587 TTCGAGCTGGACCTTCCAACTTGGAGGTGCTCCCTGGACAAACACACCCCTCAAT 3646

QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
DB 3647 GTGAAAGCGGAATACTACAAAGGACAGGGGAGAACCACTTCTGACCCAGCAGCCT 3706

RESULT 13
US-09-800-198-62
; Sequence 62, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 62
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-62

Alignment Scores:
Pred. No.: 9,47e-51 Length: 8797
Score: 417.00 Matches: 72
Percent Similarity: 88.00% Conservative: 16
```

```
Db      3527 GAATATGACACCTGCCCGAGTCTCATCTGTGGGAGAAAGGACAGACCCCTGCTTCAGGGA 3586
QY      62  TyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeuAsn 81
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      3587 TTCGAGCTGGACCCCTTCCAACCTTGGAGGCTGTCCCTGGACAAACACACACACCTCAAT 3646
QY      82  IleGlnSerGlyTleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      3647 GTGAAAGCGGAATACTACAAAGGAGCAGGGGAGAACCAAGTTCCTGACCCAGCAGCCT 3706

RESULT 15
US-09-808-602-76
; Sequence 76, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 6560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-808-602-76

Alignment Scores:
Pred. No.:      8.82e-51      Length:      6560
Score:          416.00      Matches:      72
Percent Similarity: 88.00%      Conservative: 16
Best Local Similarity: 72.00%      Mismatches:  12
Query Match:    75.77%      Indels:      0
DB:              9      Gaps:      0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-808-602-76 (1-6560)
QY      2  LeuPheArgLysTrpPheAlaAlaAlaProAspLeuSerTyrTyrPheIleTyrAspLys 21
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      215 CTCTTCCAGAAAGTCATTCACAGGTTCTCCCAACCTGGCCCTACACCTTCATCTGGGACAAG 274
QY      22  ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      275 ACAGATCGGTATGGCCAAAGGGTGTATGGACTCTCAGATGCTGTGTGTCTGTCTGCGGGTTT 334
QY      42  GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGlnGly 61
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      335 GAATATGAGACCTGTCCAGCTCTAATCTCTGGGAGAAAGGACAGCCCTCCCTTCAGGGA 394
QY      62  TyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeuAsn 81
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      395 TTCGAGTGGACCCCTCCACCTCGGTGGTGGTGGTCCCTAGACAAACACCACTCCTCAAT 454
QY      82  IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      455 GTTAAAGTGGATCTCTACAAAGGCACTGGGGGAAACCAAGTTCCTGACCCAGCAGCCT 514
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 29.8807 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_1100_1200
Perfect score: 549
Sequence: 1 RLFRKFAAPDLSYFIWD.....IQSGILHKGNGENQFVSQOP 101

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418
(without alignments)

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DBV=xlh
-O=/cgn2_1/USPTO spoel/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=issued_Patents_NA -QFWT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn 1 1 258 @runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEVTIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCBUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|---------------------|-------------------|
| 1 | 73 | 13.3 | 3023 | US-09-203-453-4 | Sequence 4, Appli |
| 2 | 73 | 13.3 | 3023 | US-09-900-236-4 | Sequence 4, Appli |
| 3 | 69.5 | 12.7 | 1875 | US-08-683-743-3 | Sequence 3, Appli |
| 4 | 69.5 | 12.7 | 2367 | US-09-489-039A-5660 | Sequence 5660, Ap |
| 5 | 69.5 | 12.7 | 3532 | US-09-221-017B-799 | Sequence 799, App |
| 6 | 68.5 | 12.5 | 3852 | US-08-306-546C-1 | Sequence 1, Appli |
| 7 | 68.5 | 12.5 | 3852 | US-08-530-524A-1 | Sequence 1, Appli |
| 8 | 66.5 | 12.1 | 4403765 | 3 US-09-103-840A-2 | Sequence 2, Appli |
| 9 | 66.5 | 12.1 | 4403765 | 3 US-09-103-840A-2 | Sequence 3, Appli |
| 10 | 66 | 12.0 | 898 | 1 US-08-419-078-3 | Sequence 3, Appli |
| 11 | 66 | 12.0 | 898 | 1 US-08-726-883-3 | Sequence 3, Appli |
| 12 | 66 | 12.0 | 1322 | 1 US-08-419-078-1 | Sequence 1, Appli |

| | | | | | | |
|------|------|------|---------|---|---------------------|--------------------|
| 13 | 66 | 12.0 | 1322 | 1 | US-08-726-883-1 | Sequence 1, Appli |
| 14 | 66 | 12.0 | 1323 | 4 | US-09-023-655-55 | Sequence 55, Appl |
| 15 | 66 | 12.0 | 2936 | 4 | US-09-976-594-1044 | Sequence 1044, Ap |
| 16 | 66 | 12.0 | 3293 | 1 | US-08-030-096-1 | Sequence 1, Appli |
| c 17 | 65.5 | 11.9 | 552 | 4 | US-09-134-000C-2495 | Sequence 2495, Ap |
| c 18 | 65.5 | 11.9 | 732 | 6 | 5273901-2 | Patent No. 5273901 |
| c 19 | 65.5 | 11.9 | 732 | 6 | 5482709-1 | Patent No. 5482709 |
| 20 | 65.5 | 11.9 | 984 | 4 | US-09-894-844-89 | Sequence 89, Appl |
| 21 | 65.5 | 11.9 | 1358 | 2 | US-08-160-524A-3 | Sequence 3, Appli |
| 22 | 65.5 | 11.9 | 1559 | 2 | US-08-160-524A-1 | Sequence 1, Appli |
| 23 | 65.5 | 11.9 | 1886 | 1 | US-08-461-773-15 | Sequence 15, Appl |
| 24 | 65.5 | 11.9 | 4411529 | 3 | US-09-103-840A-1 | Sequence 1, Appli |
| c 25 | 65.5 | 11.7 | 523 | 2 | US-08-508-786-8 | Sequence 8, Appli |
| c 26 | 64.5 | 11.7 | 523 | 5 | PCT-US96-12158-8 | Sequence 7, Appli |
| c 27 | 64.5 | 11.7 | 574 | 2 | US-08-508-786-7 | Sequence 7, Appli |
| c 28 | 64.5 | 11.7 | 574 | 5 | PCT-US96-12158-7 | Sequence 6, Appli |
| c 30 | 64.5 | 11.7 | 722 | 2 | US-08-508-786-6 | Sequence 6, Appli |
| c 31 | 64.5 | 11.7 | 722 | 5 | PCT-US96-12158-6 | Sequence 5, Appli |
| c 32 | 64.5 | 11.7 | 1030 | 2 | US-08-508-786-5 | Sequence 5, Appli |
| c 33 | 64.5 | 11.7 | 1030 | 5 | PCT-US96-12158-5 | Sequence 4, Appli |
| c 34 | 64.5 | 11.7 | 1294 | 5 | PCT-US96-12158-4 | Sequence 4, Appli |
| c 35 | 64.5 | 11.7 | 1294 | 5 | PCT-US96-12158-4 | Sequence 3, Appli |
| c 36 | 64.5 | 11.7 | 1372 | 2 | US-08-508-786-3 | Sequence 3, Appli |
| c 37 | 64.5 | 11.7 | 1372 | 5 | PCT-US96-12158-3 | Sequence 2, Appli |
| c 38 | 64.5 | 11.7 | 1388 | 2 | US-08-508-786-2 | Sequence 2, Appli |
| c 39 | 64.5 | 11.7 | 1988 | 5 | PCT-US96-12158-2 | Sequence 1, Appli |
| c 40 | 64.5 | 11.7 | 2010 | 2 | US-08-508-786-1 | Sequence 1, Appli |
| c 41 | 64.5 | 11.7 | 2010 | 5 | PCT-US96-12158-1 | Sequence 1, Appli |
| c 42 | 64.5 | 11.7 | 3470 | 4 | US-09-976-594-345 | Sequence 345, App |
| 43 | 64 | 11.7 | 2610 | 2 | US-08-989-386-2 | Sequence 2, Appli |
| 44 | 64 | 11.7 | 3045 | 4 | US-09-134-000C-2860 | Sequence 2860, Ap |
| 45 | 63.5 | 11.6 | 1724 | 3 | US-08-993-260-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-09-203-453-4
; Sequence 4, Application US/09203453
; Patent No. 6426411
; GENERAL INFORMATION:
; APPLICANT: Spiegelman, Bruce M.; Puigserver, Pere; Wu, Zhidan and Adelman, Guillaume
; TITLE OF INVENTION: PGC-1, A NOVEL BROWN FAT PPAR[SYMBOL 103 \f "Symbol"] COACTIVATOR
; FILE REFERENCE: DFN-023CP
; CURRENT APPLICATION NUMBER: US/09/203,453
; CURRENT FILING DATE: 1998-12-01
; EARLIER APPLICATION NUMBER: 09/085,912
; EARLIER FILING DATE: 1998-05-29
; EARLIER APPLICATION NUMBER: 60/048,107
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 3023
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)..(2482)
US-09-203-453-4
Alignment Scores: 2.47 Length: 3023
Pred. No.: 73.00 Matches: 29
Score: 41.58% Conservative: 13
Percent Similarity: 28.71% Mismatches: 39
Best Local Similarity: 13.30% Indels: 20
Query Match: 13.30% Gaps: 5
DB: 4
US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-203-453-4 (1-3023)
QY 3 PheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLysThr 22

| | | | |
|----------------|-----|---|-----|
| D _b | 221 | GACTTGGATACAGACAGCTTTCTGGGTGCACTCAAGTGCGTCAGTGACCACATCAGAATA | 280 |
| | | : : : : : : | |
| Q _y | 80 | LeuAsnIleGln-----SerGlyIleLeuHisLysGlyAsnGlyGluAsn | 94 |
| | | : : : : : : | |
| D _b | 281 | ATATCCAATCATGTCACAACATGAGCCCTTCAAACATATTGTGAAGAGATAGATGAAGAGAAT | 340 |
| | | : : : : : : | |
| Q _y | 95 | Gln | 95 |
| | | : : : | |
| D _b | 341 | GAG | 343 |

```

RESULT 3
US-08-683-743-3
; Sequence 3, Application US/08683743
; Patent No. 5843697
; GENERAL INFORMATION:
; APPLICANT: Pestka, Sidney
; APPLICANT: Kostenko, Serguei
; TITLE OF INVENTION: CYTOKINE RECEPTOR SIGNAL TRANSDUCTION
; TITLE OF INVENTION: CHAIN
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
;

```

ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza,
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,743
FILING DATE: 17-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-050
TELECOMMUNICATION INFORMATION:

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-08-683-743-3

```

| | | | |
|------------------------|--------|---------------|----|
| Score: | 69.50 | Matches: | 20 |
| Percent Similarity: | 45.21% | Conservative: | 13 |
| Best Local Similarity: | 27.40% | Mismatches: | 31 |
| Query Match: | 12.66% | Indels: | 9 |
| DB: | 2 | Gaps: | 3 |

US-10-029-020-14_C0BY_1100_1200 (1-101) x US-08-683-743-3 (1-1875)

```
QY      10 AlaProAspLeu---SertYrTyrPheIleTirAspLysThrAspValTyrAsnGlnLys 28
|||||   ::::   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db      440 GCCCTAAATTCAGAAATGAATACGAAACTTGCACTATGAAGATGTGTATTAATCATGG 499

QY      29 ValPhe-----GlyLeuSerGluAlaPhe---ValSerValGly 40
|||   ::::   |||   |||   |||   |||   |||   |||   |||   |||
Db      500 ACTTATATGTGCATPACTGAAAAACGGTACTGATGAAAGTTTCAAAATTAATCTCCAC 559
```

| | | |
|----|---|-----|
| 41 | TyrGluTyrGluSerCysProAspLeulleuLeuTrpGluLysArgfhrThrValLeuGln | 60 |
| QY | | |
| DB | TATGACTTTGAGGTCTCAGAACTCGAGCATCGACAATATTGTTGTTCAAGTTGGA | 619 |
| QY | 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSer | 73 |
| DB | CGGTTCTCTCTCGATCGGAAACAAAGCTGGGAAATGGAGT | 658 |

RESULT 4

```

RESOLUTION 1
US-09-489-039A-5660
; Sequence 5660, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5660
; LENGTH: 2367
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5660

```

| | | |
|------------------------|--------|------------------|
| Alignment Scores: | | |
| Pred. No.: | 5.63 | Length: 2367 |
| Score: | 69.50 | Matches: 26 |
| Percent Similarity: | 36.97% | Conservative: 18 |
| Best Local Similarity: | 21.95% | Mismatches: 36 |
| Query Match: | 12.66% | Indels: 39 |
| DR: | 4 | Gaps: 6 |

US-10-029-020-14 COPY 1100 1200 (1-101) X US-09-489-039A-5660 (1-2367)

| | | | | |
|----|--|-----|---|------|
| Qy | | 3 | PheArgIysTrpPheAlaAalaProAspLeuSerTyrTyrrPheIleTrpAspIysThr | 22 |
| Db | | 676 | TATCGCGCTGCGAGCGGAACGGCAGCATCTGATTATTACTGTTTACC GGTAAG--- | 732 |
| Qy | | 23 | AspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheVal- | 37 |
| Db | | 733 | ---CAGTGCTGGATGTGACCAAAAGCCTTCGTCGCCTGACCCGTAAA | 777 |
| Qy | | 38 | -----SerValGlyTyr-----GluTyrGluSer | 45 |
| Db | | 778 | ACGCTGTTTGGACCAAATGAGCCTCGCTCACGCGGCTCCACCATGCACATATACCGAC | 837 |
| Qy | | 46 | CysProAsp-----LeuIleLeuTrpGluLysArgThrThr | 57 |
| Db | | 838 | GCGCGGATGCGCAAAATCAGCTGATGAATTTATCCGCCCTGTGCGAGCAGCAGCGGATC | 897 |
| Qy | | 58 | VallLeuGlnGlyTyrGluIleAspla-----SerLysLeuGlyGly- | 71 |
| Db | | 898 | CCCTGCGATTGTTCCAGCTCTCTTCCGGCTATACCTCAATCAACGCAAGCGCTACGTC | 957 |
| Qy | | 72 | -----TrpSerLeuAspIysHisAlaLeuAsnIleGlnSerGlyIleLeuHis | 88 |
| Db | | 958 | TTTAACCTGCAACTGACAAGGTGCGCAGCCCCAAGATGATGAGCAGCTGTTTCA | 1014 |

RESULT 5

RESULT 5
US-09-221-017B-799/c
; Sequence 799, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: ROSS, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD

```

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSQ for Windows Version 2.0.b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 799:
SEQUENCE CHARACTERISTICS:
LENGTH: 3532 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...3532

```

| | | |
|------------------------|--------|------------------|
| Alignment Scores: | | |
| Pred. No.: | 10 | Length: 3532 |
| Score: | 69.50 | Matches: 23 |
| Percent Similarity: | 37.89% | Conservative: 13 |
| Best Local Similarity: | 24.21% | Mismatches: 32 |
| Query Match: | 12.66% | Indels: 27 |
| DR: | 4 | Gaps: 3 |

```

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-221-017B-799 (1-3532)

Qy      32 LeuSerCluAlaPheValSerValGlyTyr-----GluTyrGlu  44
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2705 CTCGATCAATTTTCGATGATGAAGGCTACTACCCGCATTGCCACAGGACCAAGGAA  2646

Qy      45 SerCysProAspLeuLeuLeuTlpGluLyArgThrThrVal-----  58
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2645 ACCGACCTACCATACACTCTTGGAGACGGCAGCAGTAGCCGTTCGTCGTCTATT  2586

Qy      59 -----LeuGlnGlyTyrGluLeAspAla  66
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2585 GCATTACGATTATTGGTGCCTTAGCCGATTGGCACAGTTCTCCGCCACAGCAGCAAGCG  2526

Qy      67 SerIysLeuGlyGlyTlpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIle  86
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```



```
Db 1686 AGTGACTGCTGGAATTGGAGAACTTTCCCAAGCTCGA---TTTGATATCT---GAGTAT 1739
QY 42 GluTyrGluSerCysProAspLeuLeuTyrGluLysArgThrThrValLeuGlnGly 61
Db 1740 GGATATCAGTCTGGCCCTTCCTCAGTACATTAGAAAGGTTTCCTCT-----1787
QY 62 TyrGluLeuAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeuAsn 81
Db 1788 ---GAAGAGCAGTGTCTTACAGAACGAGCTTTGCACITTCATCGGCAACATTGTGATTAC 1844
QY 82 IleGlnSerGlyIleLeuHisLys 89
Db 1845 GGTAAACATGAATGCTTCCACGAG 1868

RESULT 8
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 6.2e+05 Length: 4403765
Score: 66.50 Matches: 22
Percent Similarity: 36.27% Conservative: 15
Best Local Similarity: 21.57% Mismatches: 34
Query Match: 12.11% Indels: 31
DB: 3 Gaps: 4

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-103-840A-2 (1-4403765)
QY 4 ArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp-----20
Db 3372455 CGGCTGTGGGTAGCA-----GACCTCACCTATGTGTGCGACCTGGCGAGGTTTCGCC 3372505
QY 21 -----LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPhe 36
Db 3372506 TACGTGGCCTTTGTCCAGCGCCTACGCTCGAGGATCTCGGC-----1985959
QY 37 ValSerValGlyTyrGluTyrGluSerCysProAspLeuLeuTyrGluLysArgThr 56
Db 3372551 -----TGGCGGGTCGCTTCC 3372565
QY 57 ThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLys 76
Db 3372566 ACGATGCCACCTCCATGGTCTCTCGACGCGATCGAGCAAGCATCTGACCCGCCAACAA 3372625
QY 77 HisHisAlaLeuAsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPhe 96
Db 3372626 GAAAGCGTACTCGACCTGAAAGAGCTTATCCACCATACGGATAGGGA---TCTCAGTAC 3372682
QY 97 ValSer 98
Db 3372683 ACATCG 3372688
```

```
RESULT 9
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 6.2e+05 Length: 4403765
Score: 66.50 Matches: 22
Percent Similarity: 36.27% Conservative: 15
Best Local Similarity: 21.57% Mismatches: 34
Query Match: 12.11% Indels: 31
DB: 3 Gaps: 4

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-103-840A-2 (1-4403765)
QY 4 ArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp-----20
Db 1986054 CGGCTGTGGGTAGCA-----GACCTCACCTATGTGTGCGACCTGGCGAGGTTTCGCC 1986004
QY 21 -----LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPhe 36
Db 1986003 TACGTGGCCTTTGTCCAGCGCCTACGCTCGAGGATCTCGGC-----1985959
QY 37 ValSerValGlyTyrGluTyrGluSerCysProAspLeuLeuTyrGluLysArgThr 56
Db 1985958 -----TGGCGGGTCGCTTCC 1985944
QY 57 ThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLys 76
Db 1985943 ACGATGCCACCTCCATGGTCTCTCGACGCGATCGAGCAAGCATCTGACCCGCCAACAA 1985884
QY 77 HisHisAlaLeuAsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPhe 96
Db 1985883 GAAAGCGTACTCGACCTGAAAGAGCTTATCCACCATACGGATAGGGA---TCTCAGTAC 1985827
QY 97 ValSer 98
Db 1985826 ACATCG 1985821

RESULT 10
US-08-419-078-3
; Sequence 3, Application US/08419078
; Patent No. 5587306
; GENERAL INFORMATION:
; APPLICANT: HAWKINS, PHILLIP R.
; APPLICANT: SEILHAMER, JEFFREY J.
; TITLE OF INVENTION: PHOSPHOLIPASE C HOMOLOG
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3330 HILLVIEW AVENUE
```

```

; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,078
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF0030 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-855-0572
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: Human Mast Cell
; CLONE: 9118
; US-08-419-078-3

Alignment Scores:
Pred. No.: 4.48 Length: 898
Score: 66.00 Matches: 14
Percent Similarity: 54.29% Conservative: 5
Best Local Similarity: 40.00% Mismatches: 12
Query Match: 12.02% Indels: 4
DB: 1 Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x US-08-419-078-3 (1-898)
Qy 52 TrpGluLysArgThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGly 71
Db 91 TGGAGAAAGCAGCAACTGCATTAAA-----GATGTTGTCAAAGTTGGTGCA 138
Qy 72 TrpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIle 86
Db 139 GTTGATGCAGATAGCATATCCCTAGGAGTCAGTATGTTGTT 183

RESULT 11
US-08-726-883-3
; Sequence 3, Application US/08726883
; Patent No. 5676946
; GENERAL INFORMATION:
; APPLICANT: HAWKINS, PHILLIP R.
; APPLICANT: SEILHAMER, JEFFREY J.
; TITLE OF INVENTION: PHOSPHOLIPASE C HOMOLOG
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3330 HILLVIEW AVENUE
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,883

; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,078
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF0030 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-855-0572
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: Human Mast Cell
; CLONE: 9118
; US-08-726-883-3

Alignment Scores:
Pred. No.: 4.48 Length: 898
Score: 66.00 Matches: 14
Percent Similarity: 54.29% Conservative: 5
Best Local Similarity: 40.00% Mismatches: 12
Query Match: 12.02% Indels: 4
DB: 1 Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x US-08-726-883-3 (1-898)
Qy 52 TrpGluLysArgThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGly 71
Db 91 TGGAGAAAGCAGCAACTGCATTAAA-----GATGTTGTCAAAGTTGGTGCA 138
Qy 72 TrpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIle 86
Db 139 GTTGATGCAGATAGCATATCCCTAGGAGTCAGTATGTTGTT 183

RESULT 12
US-08-419-078-1
; Sequence 1, Application US/08419078
; Patent No. 5587306
; GENERAL INFORMATION:
; APPLICANT: HAWKINS, PHILLIP R.
; APPLICANT: SEILHAMER, JEFFREY J.
; TITLE OF INVENTION: PHOSPHOLIPASE C HOMOLOG
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3330 HILLVIEW AVENUE
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,078
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF0030 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
```

```
; TELEFAX: 415-855-0572
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1322 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: NO. 5597306e
; CLONE: 9118
; US-08-419-078-1

Alignment Scores:
Pred. No.: 7.84 Length: 1322
Score: 66.00 Matches: 14
Percent Similarity: 54.29% Conservative: 5
Best Local Similarity: 40.00% Mismatches: 12
Query Match: 12.02% Indels: 4
DB: 1 Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x US-08-419-078-1 (1-1322)
QY 52 TtpGluLysArgThrThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGly 71
Db 193 TGAAGAAGAGCAGCACTGCATTAAAA-----GATGTTGTCAAAGTTGGTGCA 240
QY 72 TtpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIle 86
Db 241 GTTGATGCAGATAGCATCATCTCCCTAGGAGTCAGTATGGTGT 285

RESULT 13
; Sequence 1, Application US/08726883
; Patent No. 5676946
; GENERAL INFORMATION:
; APPLICANT: HAWKINS, PHILLIP R.
; APPLICANT: SEILHAMER, JEFFREY J.
; TITLE OF INVENTION: PHOSPHOLIPASE C HOMOLOG
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3330 HILLVIEW AVENUE
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,883
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/419,078
; FILING DATE: 10-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF0030 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-855-0572
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1322 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
```

```
; IMMEDIATE SOURCE:
; LIBRARY: NO. 5676946e
; CLONE: 9118
; US-08-726-883-1

Alignment Scores:
Pred. No.: 7.84 Length: 1322
Score: 66.00 Matches: 14
Percent Similarity: 54.29% Conservative: 5
Best Local Similarity: 40.00% Mismatches: 12
Query Match: 12.02% Indels: 4
DB: 1 Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x US-08-726-883-1 (1-1322)
QY 52 TtpGluLysArgThrThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGly 71
Db 193 TGAAGAAGAGCAGCACTGCATTAAAA-----GATGTTGTCAAAGTTGGTGCA 240
QY 72 TtpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIle 86
Db 241 GTTGATGCAGATAGCATCATCTCCCTAGGAGTCAGTATGGTGT 285

RESULT 14
; Sequence 55, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: COCKS, Benjamin G.
; APPLICANT: SUSAN G. STUART
; APPLICANT: JEFFREY J. SEILHAMER
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1323 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 009118
; US-09-023-655-55

Alignment Scores:
```

Mon Aug 16 09:01:05 2004

```

Pred. No.: 7.85 Length: 1323
Score: 66.00 Matches: 14
Percent Similarity: 54.29% Conservative: 5
Best Local Similarity: 40.00% Mismatches: 12
Query Match: 12.02% Indels: 4
DB: Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-023-655-55 (1-1323)
Qy 52 TrpGluLysArgThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGly 71
Db 193 TGGAGAAAGCAGCAACTGCATTAA-----GATGTTGCAAGTTGGTGCA 240
Qy 72 TrpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIle 86
Db 241 GTTGATGCAGTAAGCATCATTCCTAGGAGTCAGTATGTGT 285

RESULT 15
US-09-976-594-1044
; Sequence 1044, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; PRIORITY FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 1044
; LENGTH: 2936
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 1383263.14
US-09-976-594-1044

Alignment Scores:
Pred. No.: 24.9 Length: 2936
Score: 66.00 Matches: 14
Percent Similarity: 54.29% Conservative: 5
Best Local Similarity: 40.00% Mismatches: 12
Query Match: 12.02% Indels: 4
DB: Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-976-594-1044 (1-2936)
Qy 52 TrpGluLysArgThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGly 71
Db 299 TGGAGAAAGCAGCAACTGCATTAA-----GATGTTGCAAGTTGGTGCA 346
Qy 72 TrpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIle 86
Db 347 GTTGATGCAGTAAGCATCATTCCTAGGAGTCAGTATGTGT 391

Search completed: August 14, 2004, 20:09:55
Job time : 1824.88 secs
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 1184.51 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_1100_1200
Perfect score: 549
Sequence: 1 RLFRKWAAPDLVSYYFWD.....IQSGILHKGNGENQFVSQOP 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QPMI=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020.cgn_1_13135_@runat_06082004_112216_29287 -NCFU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 549 | 100.0 | 6246 | 29 | AY413475 | AY413475 Homo sapi |
| 2 | 544 | 99.1 | 3190 | 29 | AY413477 | AY413477 Mus muscu |
| 3 | 544 | 99.1 | 5970 | 29 | AY413476 | AY413476 Pan trogl |
| 4 | 454 | 82.7 | 5087 | 29 | AY405422 | AY405422 Mus muscu |
| 5 | 453 | 82.5 | 5069 | 29 | AY405421 | AY405421 Pan trogl |
| 6 | 453 | 82.5 | 5094 | 29 | AY405420 | AY405420 Homo sapi |
| 7 | 422 | 76.9 | 796 | 14 | CD578516 | CD578516 UI-M-G10- |
| 8 | 403 | 73.4 | 750 | 9 | AL710250 | AL710250 DKFZp686F |
| 9 | 391 | 71.2 | 539 | 9 | AL919858 | AL919858 AL919858 |
| 10 | 390 | 71.0 | 725 | 14 | CA752216 | CA752216 UI-M-FOO- |
| 11 | 384 | 69.9 | 789 | 14 | CD803315 | CD803315 UI-M-GV0- |
| 12 | 376 | 68.5 | 348 | 14 | CB704746 | CB704746 AMGNNUC:S |
| 13 | 348 | 63.4 | 262 | 10 | BB538418 | BB538418 BB538418 |
| 14 | 290 | 52.8 | 318 | 10 | BB501284 | BB501284 BB501284 |
| 15 | 287 | 52.3 | 505 | 12 | BJ016612 | BJ016612 BJ016612 |
| 16 | 268 | 48.8 | 728 | 14 | CF536185 | CF536185 UI-M-G10- |
| 17 | 248.5 | 45.3 | 919 | 29 | CNS01K9E | AL148051 Anopheles |
| 18 | 223 | 40.6 | 692 | 14 | CF530821 | CF530821 UI-M-GH0- |
| 19 | 223 | 40.6 | 706 | 29 | CE611069 | CE611069 tigr-gss- |
| 20 | 222 | 40.4 | 529 | 9 | AL699161 | AL699161 DKFZp686F |
| 21 | 218 | 39.7 | 716 | 29 | AG098731 | AG098731 Pan trogl |
| 22 | 215.5 | 39.3 | 936 | 29 | CNS02ZKC | AL221061 Tetraodon |
| 23 | 210 | 38.3 | 833 | 29 | CNS02D8V | AL192136 Tetraodon |
| 24 | 208.5 | 38.0 | 346 | 28 | AQ970978 | AQ970978 RPCI-23-3 |
| 25 | 203 | 37.0 | 778 | 13 | BU705631 | BU705631 UI-M-FO0- |
| 26 | 201 | 36.6 | 736 | 14 | CD803545 | CD803545 UI-M-GV0- |
| 27 | 170 | 31.0 | 730 | 14 | CF182768 | CF182768 UI-M-EY0- |
| 28 | 161.5 | 29.4 | 1146 | 28 | CC238800 | CC238800 CH261-50M |
| 29 | 156 | 28.4 | 835 | 28 | BZ165653 | BZ165653 CH230-463 |
| 30 | 141 | 25.7 | 851 | 29 | CNS03J5I | AL246447 Tetraodon |
| 31 | 141 | 25.7 | 990 | 29 | CNS0428B | AL271172 Tetraodon |
| 32 | 134 | 24.4 | 975 | 29 | CNS05RSP | AL350962 Tetraodon |
| 33 | 132 | 24.0 | 619 | 29 | CI22E8 | AL226991 Ciona int |
| 34 | 124 | 22.6 | 480 | 9 | AI387033 | AI387033 GH17448.5 |
| 35 | 123 | 22.4 | 735 | 14 | CF745263 | CF745263 UI-M-GV0- |
| 36 | 122.5 | 22.3 | 968 | 29 | CNS00765 | AL066865 Drosophil |
| 37 | 107 | 19.5 | 726 | 29 | CC574461 | CC574461 CH240-451 |
| 38 | 106 | 19.3 | 1052 | 29 | CNS05D17 | AL331828 Tetraodon |
| 39 | 106 | 19.3 | 1085 | 29 | CNS0533P | AL318958 Tetraodon |
| 40 | 105 | 19.1 | 1201 | 29 | CNS0109M | AL098692 Drosophil |
| 41 | 100 | 18.2 | 916 | 29 | CNS04M38 | AL296909 Tetraodon |
| 42 | 94 | 17.1 | 243 | 28 | AZ647907 | AZ647907 IM0514K07 |
| 43 | 84 | 15.3 | 645 | 28 | BH073129 | BH073129 RPCI-24-2 |
| 44 | 84 | 15.3 | 726 | 29 | CC611675 | CC611675 OGUAC85TM |
| 45 | 84 | 15.3 | 733 | 29 | CE320811 | CE320811 tigr-gss- |

ALIGNMENTS

RESULT 1
AY413475
LOCUS
DEFINITION Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY413475
VERSION AY413475.1 GI:39769437
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 6246)

AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 6246)

AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..6246

gene /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>6246
/locus_tag="HCM4903"

ORIGIN

Alignment Scores:
Pred. No.: 1.38e-63 Length: 6246
Score: 549.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 29 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AY413475 (1-6246)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 1249 CGCCTCTTCAGGAAGTGGTTGGTGGAGCTGCTGAGCTGCTTATTTCATTTGGGAC 1308

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 1309 AAGACAGAGCTTACAAACAGAGGTTTGGGCTTTCAGAGCCTTGTTCGTTGGGT 1368

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrValLeuGln 60
Db 1369 TATGAATATGAATCTCCAGAGTCTTAATCTCTGGGAAAGAAAGAACACAGTGTGCAG 1428

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db 1429 GGCATGAATTCAGCGGTCCAGCTGGAGATGGAGCTAGACAAACATCATGCCCTC 1488

QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 1489 AACATTCAAAGTGGCATCTCTGCACAAAGGGAATGGGAGAACACCATGTTGTCTCAGCAG 1548

QY 101 Pro 101
Db 1549 CCT 1551

RESULT 2
AY413477
LOCUS Mus musculus HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY413477
VERSION AY413477.1 GI:39769439
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 3190)

AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 3190)

AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..3190

gene /organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>3190
/locus_tag="HCM4903"

ORIGIN

Alignment Scores:
Pred. No.: 2.57e-63 Length: 3190
Score: 544.00 Matches: 100
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 1
Query Match: 99.09% Indels: 0
DB: 29 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AY413477 (1-3190)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 1249 CGCCTCTTCAGGAAGTGGTTGGTGGAGCTGCTGAGCTGCTTATTTCATTTGGGAC 1308

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 1309 AAGACAGAGTCTACAAACAGAGGTTTGGGCTTTCAGAGCCTTGTTCGTTGGGT 1368

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrValLeuGln 60
Db 1369 TATGAATATGAATCTCCAGAGTCTTAATCTCTGGGAAAGAAAGAACACAGTGTGCAG 1428

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db 1429 GGCATGAATTCAGCGGTCCAGCTGGAGATGGAGCTAGACAAACATCATGCCCTC 1488

QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 1489 AACATCCAGAGTGGCATCTCTGCACAAAGGGAATGGGAGAACACCATGTTGTCTCCACAG 1548

QY 101 Pro 101
Db 1549 CCA 1551

RESULT 3
AY413476
LOCUS Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY413476
VERSION AY413476.1 GI:39769438
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
ORIGIN

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
1 (bases 1 to 5970)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Infering nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 5970)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..5970
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>5970
/locus_tag="HCM4903"

Alignment Scores:
Pred. No.: 6.32e-63 Length: 5970
Score: 544.00 Matches: 100
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 1
Query Match: 99.09% Indels: 0
DB: 29 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AY413476 (1-5970)

Qy 1 ArgLeuPheArgLysTyrPheAlaAlaProAspLeuSerTyrTyrPheIleTrrpAsp 20
Db 973 CGCCTCTTCAGGAAGTGGTTCGTCGAGCCCGACATCTGTCCTATTATTCATTGGGAC 1032
Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 1033 AAGACAGACGCTACACAGAGGTGTTGGGCTTTCAGAGGCTTTGTTTCGCGGT 1092
Qy 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrrpGluLysArgThrValLeuGln 60
Db 1093 TATGAATATGAATCCTGCCAGATCTAATCTGTGGGAAAAGAACAGCAGTCTGCAG 1152
Qy 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrrpSerLeuAspLysHisAlaLeu 80
Db 1153 GGCTATGAATGTATGCTCAAGCTTGAGATGGAGCCCTAGACAAACATCATGCCCTC 1212
Qy 81 AsnIleGlnSerGlyLeuLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 1213 AACATCAAGTGCATCCTGCACAAAGGAATGGGAGACCACTTTGTCTCTCAGCAG 1272
Qy 101 Pro 101
Db 1273 CCT 1275

RESULT 4
AY405422 5087 bp DNA linear GSS 12-DEC-2003
LOCUS
DEFINITION Mus musculus HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY405422
VERSION
AY405422.1 GI:39761396
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
ORIGIN

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 5087)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Infering nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 5087)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..5087
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>5087
/locus_tag="HCM2218"

Alignment Scores:
Pred. No.: 1.32e-50 Length: 5087
Score: 454.00 Matches: 79
Percent Similarity: 91.09% Conservative: 13
Best Local Similarity: 78.22% Mismatches: 9
Query Match: 82.70% Indels: 0
DB: 29 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AY405422 (1-5087)

Qy 1 ArgLeuPheArgLysTyrPheAlaAlaProAspLeuSerTyrTyrPheIleTrrpAsp 20
Db 421 AGACTCTTCAGGAAGTGGTTCCTGCTCGCCAAACTTGGCCTACACGTTTCATCTGGGAT 480
Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 481 AAGACGAGCGCATATAATCAAGAACTCTACGGCTTGTACAGAGGCGATTGTCTCGTCGGA 540
Qy 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrrpGluLysArgThrValLeuGln 60
Db 541 TACAGATACAGTCTGCTGCTGACCTCTCTGGGAAAAGAGAGACTGCCGTTTTCGA 600
Qy 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrrpSerLeuAspLysHisAlaLeu 80
Db 601 GGCTATGAGTTGGATGCTTCGAACATGGCGCGCTGACGTTGGACAAGCACCATTGACTG 660
Qy 81 AsnIleGlnSerGlyLeuLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 661 GACCTTCAGACGCTATCTACTATACAAAGAAATGGAGAAATCAGTTCTCTCTCAGCAG 720
Qy 101 Pro 101
Db 721 CCT 723

RESULT 5
AY405421 5069 bp DNA linear GSS 12-DEC-2003
LOCUS
DEFINITION Pan troglodytes HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY405421
VERSION
AY405421.1 GI:39761395
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)

```
ORGANISM      Pan troglodytes
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
              1 (bases 1 to 5069)
              Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
              Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
              Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
              Adams, M.D. and Cargill, M.
TITLE        Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
JOURNAL      Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE    2 (bases 1 to 5069)
AUTHORS      Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
              Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
              Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
              Adams, M.D. and Cargill, M.
TITLE        Direct Submission
JOURNAL      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
COMMENT      This sequence was made by sequencing genomic exons and ordering
              them based on alignment.
FEATURES     Location/Qualifiers
              source
                1..5069
                /organism="Pan troglodytes"
                /mol_type="genomic DNA"
                /db_xref="taxon:9598"
                <1..>5069
                /locus_tag="HCM2218"
              gene
                1..5069
                /locus_tag="HCM2218"
              ORIGIN
                Alignment Scores:
                Pred. No.:      1.8e-50      Length:      5069
                Score:          453.00      Matches:      78
                Percent Similarity: 91.0%      Conservative: 14
                Best Local Similarity: 77.23%      Mismatches:   9
                Query Match:      82.51%      Indels:       0
                DB:               29          Gaps:         0
US-10-029-020-14_COPY_1100_1200 (1-101) x AY405421 (1-5069)
QY 1 ArgLeuPheArgLysTyrPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 422 AGACTCTTCCAAAAGTGGTTTCCTGCCTCACCACAACTTGGCCATATCTTCAATGGGAT 481
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 482 AAAACAGATGCATATAATCAGAAAGTCTATGCTCTGTAAGCTGTGTGTGTCAGTTGA 541
QY 41 TyrGluTyrGluSerCysProAspLeuLeuLeuTyrGluLysArgThrThrValLeuGln 60
Db 542 TATGATGATGATGCTGTGTGGACCTGCTGTGGGAAAAGAGACTGCCATTCGCGAG 601
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyTyrPheSerLeuAspLysHisAlaLeu 80
Db 602 GCCTATGATGATGCTGTGTGGACCTGCTGTGGGAAAAGAGACTGCCATTCGCGAG 661
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 662 GATGTACAGACGGTATATCTGACAGGAAACGGGAAACCAACCACTTCTCTCCAGCAG 721
QY 101 Pro 101
Db 722 CCT 724
RESULT 6
LOCUS      AY405420
DEFINITION Homo sapiens HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
              genomic survey sequence.
ACCESSION  AY405420
VERSION     AY405420.1 GI:39761394
KEYWORDS    GSS.

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
              1 (bases 1 to 5094)
              Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
              Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
              Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
              Adams, M.D. and Cargill, M.
TITLE        Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
JOURNAL      Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE    2 (bases 1 to 5094)
AUTHORS      Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
              Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
              Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
              Adams, M.D. and Cargill, M.
TITLE        Direct Submission
JOURNAL      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
COMMENT      This sequence was made by sequencing genomic exons and ordering
              them based on alignment.
FEATURES     Location/Qualifiers
              source
                1..5094
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                <1..>5094
                /locus_tag="HCM2218"
              gene
                1..5094
                /locus_tag="HCM2218"
              ORIGIN
                Alignment Scores:
                Pred. No.:      1.81e-50      Length:      5094
                Score:          453.00      Matches:      78
                Percent Similarity: 91.09%      Conservative: 14
                Best Local Similarity: 77.23%      Mismatches:   9
                Query Match:      82.51%      Indels:       0
                DB:               29          Gaps:         0
US-10-029-020-14_COPY_1100_1200 (1-101) x AY405420 (1-5094)
QY 1 ArgLeuPheArgLysTyrPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 422 AGACTCTTCCAAAAGTGGTTTCCTGCCTCACCACAACTTGGCCATATCTTCAATGGGAT 481
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 482 AAAACAGATGCATATAATCAGAAAGTCTATGCTCTGTAAGCTGTGTGTGTCAGTTGA 541
QY 41 TyrGluTyrGluSerCysProAspLeuLeuLeuTyrGluLysArgThrThrValLeuGln 60
Db 542 TATGATGATGATGCTGTGTGGACCTGCTGTGGGAAAAGAGACTGCCATTCGCGAG 601
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyTyrPheSerLeuAspLysHisAlaLeu 80
Db 602 GCCTATGATGATGCTGTGTGGACCTGCTGTGGGAAAAGAGACTGCCATTCGCGAG 661
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 662 GATGTACAGACGGTATATCTGACAGGAAACGGGAAACCAACCACTTCTCTCCAGCAG 721
QY 101 Pro 101
Db 722 CCT 724
RESULT 7
LOCUS      CD578516
DEFINITION UI-M-GIO-c9g-e-23-0-UI-r1 NIH BMAP_G10 Mus musculus cDNA clone
              IMAGE:6852960 5', mRNA sequence.
ACCESSION  CD578516
VERSION     CD578516.1 GI:31742907
KEYWORDS    796 bp mRNA linear EST 09-JUL-2003
```

KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 796)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouseefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5,
Location/Qualifiers
1..796
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/db_xref="IMAGE:6852960"
/clone="IMAGE:6852960"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_G10"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 2,398-47 Length: 796
Score: 422.00 Matches: 78
Percent Similarity: 89.11% Conservative: 12
Best Local Similarity: 77.23% Mismatches: 11
Query Match: 76.87% Indels: 1
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x CD578516 (1-796)

Qy 1 ArgLeuPheArgLysTyrPheAlaAlaProAspLeuSerTyrTyrPheLeuTyrPasp 20
Db 464 AGACTCTCCAGAGTGGTTCCTGCTCGCAACTGGCTACACGTTTCATCTGGAT 523

Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 524 AAGACGGACGCATATAATCAGAAAGTCTACGCTTGTTCAGAGCGAGTGTGTCGTCGGA 583

Qy 41 TyrGluTyrGluSerCysProAspLeuLeuTyrGluLysArgThrThrValLeuGln 60
Db 584 TACGAGTACGAGTCTGCTGTCGACCTGACTCTCTCGGAAAGAGGACTGCCGCTTTTGAA 643

Qy 61 GlyTyrGluLeuAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db 644 GGCTATGAGTGGATGCTTCGACATGGCGGCTGGAGTGGACAAACA-CATGACTG 702

Qy 81 AsnLeuGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 703 GACGTTCAAGCGGTATCTACTATACAAAGAAATGGAGAAATCAGTTCATCTCTCAGCAG 762

Qy 101 Pro 101
Db 763 CCT 765

RESULT 8
AL710250 750 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp686F1868_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION DKFZp686F1868 5', mRNA sequence.
ACCESSION AL710250
VERSION AL710250.1 GI:19693605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 750)
AUTHORS Ottenwaelder, B., Obermaier, B., Mewes, H.W., Mewes, H.W., Weil, B. and Wiemann, S.
TITLE EST (Ottenwaelder, B., Obermaier, B., Mewes, H.W., Weil, B. and Wiemann, S.)
JOURNAL Unpublished (2001)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr 1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by MediGenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. No s1 sequence available.
This clone (DKFZp686F1868) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Lichtenberg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686F1868"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"

ORIGIN
Alignment Scores:
Pred. No.: 9,178-45 Length: 750
Score: 403.00 Matches: 72
Percent Similarity: 87.13% Conservative: 16
Best Local Similarity: 71.29% Mismatches: 12
Query Match: 73.41% Indels: 1
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AL710250 (1-750)

Qy 2 LeuPheArgLysTyrPheAlaAlaProAspLeuSerTyrTyrPheLeuTyrPaspLys 21
Db 404 CTCCTCCAGAGTCAATCCAGGCTCTCCCAACCTGGCTACACTTCATCTGGACAAG 463

Qy 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
Db 464 ACAGATGCGTATGGCCAAAGGGTGTATGGACTCTCAGATCTGTGTGTCTGCGGTTT 523

Qy 42 GluTyrGluSerCysProAspLeuLeuTyrGluLysArgThrThrValLeuGlnGly 61
Db 524 GAATATGAGACTCTGCCAGTCTTAATTTCTCTGGAGAAAGACAGCCCTCCTTCAGGGA 583

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

| | | | |
|----|-----|--|-----|
| Qy | 1 | ArgLeuPheArgLysTyrPtpheAlaAlaProAspLeuSerTyrTyrPheIleIrrpasp | 20 |
| Db | 459 | AGACTCTTCGAAAGTGGTTCCTGCCCTCGCCAACTTGGCGTACACGTTCACTGGGAT | 518 |
| Qy | 21 | LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly | 40 |
| Db | 519 | AAGACGGACGCATATAATCAGAAAGTCTACGGCTTCAGAGCGAGTTGTGTCCGTCCGA | 578 |
| Qy | 41 | TyrGluTyrGluSerCysProAspLeulleIleuTrrpGluLysArgThrThrValLeuGln | 60 |
| Db | 579 | TACGAGTACGAGTCGTGCTTGGACCTGACTCTCTCGGAAACAGAGGACTGCCGTTTTCGAA | 638 |
| Qy | 61 | GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu | 80 |
| Db | 639 | GGCTATGAGTTGGATGCTTCGAAACATGGCGCGCTGGACGCTTCGACCAACCATGTACTGT | 698 |
| Qy | 81 | AsnIleGlnSerGlyIleLeuHisLys | 89 |
| Db | 699 | GACGTTCCAGAACGGTATCTACTATACAAA | 725 |

[illegible]

| ACCESSION | CD803315 | GI:32462141 |
|-----------|------------|-------------|
| VERSION | CD803315.1 | GI:32462141 |

| | |
|----------|----------------------------|
| KEYWORDS | EST. |
| SOURCE | Mus musculus (house mouse) |

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Cra

Mammalia; Eutheria; Rodentia; Sci

REFERENCE
1 (bases 1 to 789)

AUTHORS NIH-MGC <http://mgc.nhlbi.nih.gov/>.
TITLE National Institutes of Health Ma

ILLI
JOURNAL. Unpublished (1999)
NATIONAL INSTITUTES OF HEALTH, MD

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin,

cdNA Library preparation: Dr. M.

cdna Library Arrayed by: Dr. M.

DNA Sequencing by: Dr. M. Bento

CLONE DISTRIBUTION: DISTRIBUTION
<http://genome.uiowa.edu/distribution>

This clone was contributed by the

(BMAP)

Seq primer: pYX-5.

| FEATURES | Location/Qualifiers |
|---|---|
| 1. <i>Age</i> | 1. <i>Age</i> |
| 2. <i>Gender</i> | 2. <i>Gender</i> |
| 3. <i>Marital Status</i> | 3. <i>Marital Status</i> |
| 4. <i>Education</i> | 4. <i>Education</i> |
| 5. <i>Income</i> | 5. <i>Income</i> |
| 6. <i>Occupation</i> | 6. <i>Occupation</i> |
| 7. <i>Religion</i> | 7. <i>Religion</i> |
| 8. <i>Political Affiliation</i> | 8. <i>Political Affiliation</i> |
| 9. <i>Health Status</i> | 9. <i>Health Status</i> |
| 10. <i>Travel History</i> | 10. <i>Travel History</i> |
| 11. <i>Family Size</i> | 11. <i>Family Size</i> |
| 12. <i>Home Ownership</i> | 12. <i>Home Ownership</i> |
| 13. <i>Vehicle Ownership</i> | 13. <i>Vehicle Ownership</i> |
| 14. <i>Employment Status</i> | 14. <i>Employment Status</i> |
| 15. <i>Insurance Coverage</i> | 15. <i>Insurance Coverage</i> |
| 16. <i>Charitable Contributions</i> | 16. <i>Charitable Contributions</i> |
| 17. <i>Real Estate Holdings</i> | 17. <i>Real Estate Holdings</i> |
| 18. <i>Investment Portfolio</i> | 18. <i>Investment Portfolio</i> |
| 19. <i>Debt Levels</i> | 19. <i>Debt Levels</i> |
| 20. <i>Life Insurance Policy</i> | 20. <i>Life Insurance Policy</i> |
| 21. <i>Retirement Savings</i> | 21. <i>Retirement Savings</i> |
| 22. <i>Spouse's Income</i> | 22. <i>Spouse's Income</i> |
| 23. <i>Spouse's Occupation</i> | 23. <i>Spouse's Occupation</i> |
| 24. <i>Spouse's Education</i> | 24. <i>Spouse's Education</i> |
| 25. <i>Spouse's Health Status</i> | 25. <i>Spouse's Health Status</i> |
| 26. <i>Spouse's Travel History</i> | 26. <i>Spouse's Travel History</i> |
| 27. <i>Spouse's Family Size</i> | 27. <i>Spouse's Family Size</i> |
| 28. <i>Spouse's Home Ownership</i> | 28. <i>Spouse's Home Ownership</i> |
| 29. <i>Spouse's Vehicle Ownership</i> | 29. <i>Spouse's Vehicle Ownership</i> |
| 30. <i>Spouse's Employment Status</i> | 30. <i>Spouse's Employment Status</i> |
| 31. <i>Spouse's Insurance Coverage</i> | 31. <i>Spouse's Insurance Coverage</i> |
| 32. <i>Spouse's Charitable Contributions</i> | 32. <i>Spouse's Charitable Contributions</i> |
| 33. <i>Spouse's Real Estate Holdings</i> | 33. <i>Spouse's Real Estate Holdings</i> |
| 34. <i>Spouse's Investment Portfolio</i> | 34. <i>Spouse's Investment Portfolio</i> |
| 35. <i>Spouse's Debt Levels</i> | 35. <i>Spouse's Debt Levels</i> |
| 36. <i>Spouse's Life Insurance Policy</i> | 36. <i>Spouse's Life Insurance Policy</i> |
| 37. <i>Spouse's Retirement Savings</i> | 37. <i>Spouse's Retirement Savings</i> |
| 38. <i>Spouse's Spouse's Income</i> | 38. <i>Spouse's Spouse's Income</i> |
| 39. <i>Spouse's Spouse's Occupation</i> | 39. <i>Spouse's Spouse's Occupation</i> |
| 40. <i>Spouse's Spouse's Education</i> | 40. <i>Spouse's Spouse's Education</i> |
| 41. <i>Spouse's Spouse's Health Status</i> | 41. <i>Spouse's Spouse's Health Status</i> |
| 42. <i>Spouse's Spouse's Travel History</i> | 42. <i>Spouse's Spouse's Travel History</i> |
| 43. <i>Spouse's Spouse's Family Size</i> | 43. <i>Spouse's Spouse's Family Size</i> |
| 44. <i>Spouse's Spouse's Home Ownership</i> | 44. <i>Spouse's Spouse's Home Ownership</i> |
| 45. <i>Spouse's Spouse's Vehicle Ownership</i> | 45. <i>Spouse's Spouse's Vehicle Ownership</i> |
| 46. <i>Spouse's Spouse's Employment Status</i> | 46. <i>Spouse's Spouse's Employment Status</i> |
| 47. <i>Spouse's Spouse's Insurance Coverage</i> | 47. <i>Spouse's Spouse's Insurance Coverage</i> |
| 48. <i>Spouse's Spouse's Charitable Contributions</i> | 48. <i>Spouse's Spouse's Charitable Contributions</i> |
| 49. <i>Spouse's Spouse's Real Estate Holdings</i> | 49. <i>Spouse's Spouse's Real Estate Holdings</i> |
| 50. <i>Spouse's Spouse's Investment Portfolio</i> | 50. <i>Spouse's Spouse's Investment Portfolio</i> |
| 51. <i>Spouse's Spouse's Debt Levels</i> | 51. <i>Spouse's Spouse's Debt Levels</i> |
| 52. <i>Spouse's Spouse's Life Insurance Policy</i> | 52. <i>Spouse's Spouse's Life Insurance Policy</i> |
| 53. <i>Spouse's Spouse's Retirement Savings</i> | 53. <i>Spouse's Spouse's Retirement Savings</i> |
| 54. <i>Spouse's Spouse's Spouse's Income</i> | 54. <i>Spouse's Spouse's Spouse's Income</i> |
| 55. <i>Spouse's Spouse's Spouse's Occupation</i> | 55. <i>Spouse's Spouse's Spouse's Occupation</i> |
| 56. <i>Spouse's Spouse's Spouse's Education</i> | 56. <i>Spouse's Spouse's Spouse's Education</i> |
| 57. <i>Spouse's Spouse's Spouse's Health Status</i> | 57. <i>Spouse's Spouse's Spouse's Health Status</i> |
| 58. <i>Spouse's Spouse's Spouse's Travel History</i> | 58. <i>Spouse's Spouse's Spouse's Travel History</i> |
| 59. <i>Spouse's Spouse's Spouse's Family Size</i> | 59. <i>Spouse's Spouse's Spouse's Family Size</i> |
| 60. <i>Spouse's Spouse's Spouse's Home Ownership</i> | 60. <i>Spouse's Spouse's Spouse's Home Ownership</i> |
| 61. <i>Spouse's Spouse's Spouse's Vehicle Ownership</i> | 61. <i>Spouse's Spouse's Spouse's Vehicle Ownership</i> |
| 62. <i>Spouse's Spouse's Spouse's Employment Status</i> | 62. <i>Spouse's Spouse's Spouse's Employment Status</i> |
| 63. <i>Spouse's Spouse's Spouse's Insurance Coverage</i> | 63. <i>Spouse's Spouse's Spouse's Insurance Coverage</i> |
| 64. <i>Spouse's Spouse's Spouse's Charitable Contributions</i> | 64. <i>Spouse's Spouse's Spouse's Charitable Contributions</i> |
| 65. <i>Spouse's Spouse's Spouse's Real Estate Holdings</i> | 65. <i>Spouse's Spouse's Spouse's Real Estate Holdings</i> |
| 66. <i>Spouse's Spouse's Spouse's Investment Portfolio</i> | 66. <i>Spouse's Spouse's Spouse's Investment Portfolio</i> |
| 67. <i>Spouse's Spouse's Spouse's Debt Levels</i> | 67. <i>Spouse's Spouse's Spouse's Debt Levels</i> |
| 68. <i>Spouse's Spouse's Spouse's Life Insurance Policy</i> | 68. <i>Spouse's Spouse's Spouse's Life Insurance Policy</i> |
| 69. <i>Spouse's Spouse's Spouse's Retirement Savings</i> | 69. <i>Spouse's Spouse's Spouse's Retirement Savings</i> |
| 70. <i>Spouse's Spouse's Spouse's Spouse's Income</i> | 70. <i>Spouse's Spouse's Spouse's Spouse's Income</i> |
| 71. <i>Spouse's Spouse's Spouse's Spouse's Occupation</i> | 71. <i>Spouse's Spouse's Spouse's Spouse's Occupation</i> |
| 72. <i>Spouse's Spouse's Spouse's Spouse's Education</i> | 72. <i>Spouse's Spouse's Spouse's Spouse's Education</i> |
| 73. <i>Spouse's Spouse's Spouse's Spouse's Health Status</i> | 73. <i>Spouse's Spouse's Spouse's Spouse's Health Status</i> |
| 74. <i>Spouse's Spouse's Spouse's Spouse's Travel History</i> | 74. <i>Spouse's Spouse's Spouse's Spouse's Travel History</i> |
| 75. <i>Spouse's Spouse's Spouse's Spouse's Family Size</i> | 75. <i>Spouse's Spouse's Spouse's Spouse's Family Size</i> |
| 76. <i>Spouse's Spouse's Spouse's Spouse's Home Ownership</i> | 76. <i>Spouse's Spouse's Spouse's Spouse's Home Ownership</i> |
| 77. <i>Spouse's Spouse's Spouse's Spouse's Vehicle Ownership</i> | 77. <i>Spouse's Spouse's Spouse's Spouse's Vehicle Ownership</i> |
| 78. <i>Spouse's Spouse's Spouse's Spouse's Employment Status</i> | 78. <i>Spouse's Spouse's Spouse's Spouse's Employment Status</i> |
| 79. <i>Spouse's Spouse's Spouse's Spouse's Insurance Coverage</i> | 79. <i>Spouse's Spouse's Spouse's Spouse's Insurance Coverage</i> |
| 80. <i>Spouse's Spouse's Spouse's Spouse's Charitable Contributions</i> | 80. <i>Spouse's Spouse's Spouse's Spouse's Charitable Contributions</i> |
| 81. <i>Spouse's Spouse's Spouse's Spouse's Real Estate Holdings</i> | 81. <i>Spouse's Spouse's Spouse's Spouse's Real Estate Holdings</i> |
| 82. <i>Spouse's Spouse's Spouse's Spouse's Investment Portfolio</i> | 82. <i>Spouse's</i> |

```
source
1. .789
```

```
/organism="Mus musculus"
/mol time="mpna"
```

```

/mol_type="MKN"
/strain="C57BL/6"

```

```
/db xref="taxon:10090"
```

```

/clone="IMAGE:30545857"

```

/tissue_type="whole brain

```
/dev_stage="1, 5, and 15
```

```
/lab_host="DH10B (T1 phas
```

```
/clone_lib="NIH_BMAP_GVO
/not= "Organ: Brain: Vec
```

```
/note=="Organ: Brain; Vec
```

BonaÍdo. Lennon and Soar
 site 2: Not 1; the lib

1996. Denatured RNA was

gel.First strand cDNA sy

primer containing a Not

size selected according

ECOR I adaptor, digested

directionally into pYX-A

sequence located between
is CGAAGTGAAT. This lib

Iowa Brain Anatomy Professor

Developing Mouse Nervous

ORIGIN

Alignment Scores:
 Pred. No.: 4.13e-42 Length: 789
 Score: 384.00 Matches: 67
 Percent Similarity: 81.57% Conservative: 9
 Best Local Similarity: 80.72% Mismatches: 7
 Query Match: 69.95% Indels: 0
 DB: 14 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x CD803315 (1-789)

| | | | |
|----|-----|--|-----|
| QY | 19 | TrpAspLysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSer | 38 |
| DB | 2 | TGGGATAAGACGACGCATATAATCAGAAAGTCTACGGCTTGTTCAGAGGAGTTGTGTCC | 61 |
| QY | 39 | ValGlyTyrGluTyrGluSerCysProAspLeuIleuTyrGluLysArgThrVal | 58 |
| DB | 62 | GTCGGATACAGTACGAGTCGTGTGGACCTGACTCTCTGGGAAAAGAGGACTGCCGTT | 121 |
| QY | 59 | LeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHis | 78 |
| DB | 122 | TTCGAAGCTATCAGTGTGGATGCTTCGAACATGGCGCGCTGGAGCTTGGACAAGCACCAT | 181 |
| QY | 79 | AlaLeuAsnIleGlnSerGlyIleLeuHisGlyAsnGlyGluAsnGlnPheValSer | 98 |
| DB | 182 | GTACTGACGCTTCAGAACGGTATACTATACAAAGAAATGAGAAATCAGTTTCATCTCT | 241 |
| QY | 99 | GlnGlnPro 101 | |
| DB | 242 | CAGCAGCCT 250 | |

RESULT 12
 CB704746
 LOCUS
 DEFINITION
 CB704746
 clone srp2-00285-d10 5', mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 (bases 1 to 348)
 Angen EST Program.
 AUTHORS
 TITLE
 Angen Rat EST Program
 JOURNAL
 COMMENT
 Contact: Dan Fitzpatrick
 Angen, Inc
 One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4881
 Plate: 00285 row: d column: 10.

FEATURES
 source
 1..348
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="srp2-00285-d10"
 /tissue_type="prostate tissue"
 /clone_lib="srp2 (10220)"
 /note="vector: pSPOR1; Site 1: SalI; Site 2: NotI; rat
 prostate normalized double selected poly(A+) mRNA size
 fraction > 1 kb"

ORIGIN

Alignment Scores:
 Pred. NO.: 1.62e-41 Length: 348
 Score: 376.00 Matches: 69
 Percent Similarity: 87.50% Conservative: 15


```

DB:          12          Gaps:          0
US-10-029-020-14_COPY_1100_1200 (1-101) x BJ016612 (1-505)
QY   34 GluAlaPheValSerValGlyTyrGluTyrGluSerCysProAspLeuIleLeuTrpGlu 53
Db   503 GAGGCAGTTGTGTCGGTGGGTTTGAATACGAGTCCTTGACCTCATCTTGTGGGAG 444
QY   54 LysArgThrThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSer 73
Db   443 AAGAGGACGGCCATCTTACAAGGCTACGAGATGGACGCGTCCAAACATGGGAGGATGGACG 384
QY   74 LeuAspLysHisHisAlaLeuAsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGlu 93
Db   383 TTGGACAAACATCACATCCTGGATGTGCAAGCGGTATCTTGATATAAGGGCAACGGTGAA 324
QY   94 AsnGlnPheValSerGlnGlnPro 101
Db   323 AATACTTTGTATCCCTGCTACCT 300

```

Search completed: August 14, 2004, 18:03:09
 Job time : 1197.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 171.323 Seconds
(without alignments)

2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_1100_1200

Perfect score: 549

Sequence: 1 RLFRKWFKAAPDLSYFIWD.....IQSGILHKGNGENQFVSQOP 101

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 segs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/cgn2.1/USPTO.spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N Geneseg 29Jan04 -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CNG 1.1.1868 @runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Geneseg 29Jan04.*
- 2: Geneseg 1990s.*
- 3: Geneseg 2000s.*
- 4: Geneseg 2001as.*
- 5: Geneseg 2001bs.*
- 6: Geneseg 2002s.*
- 7: Geneseg 2003as.*
- 8: Geneseg 2003bs.*
- 9: Geneseg 2003cs.*
- 10: Geneseg 2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|-------------|
| 1 | 549 | 100.0 | 8354 | 6 ABS52100 | Human TEN |
| 2 | 549 | 100.0 | 8645 | 6 ABS78652 | Human cDN |
| 3 | 538.5 | 98.1 | 8438 | 6 ABN85378 | Human NOV |
| 4 | 532 | 96.9 | 453 | 3 AAC74626 | Human ORF |
| 5 | 453 | 82.5 | 8473 | 6 ABQ82345 | Human NOV |
| 6 | 453 | 82.5 | 8487 | 6 ABQ82346 | Human NOV |
| 7 | 453 | 82.5 | 8645 | 6 ABQ82344 | Human NOV |
| 8 | 453 | 82.5 | 8675 | 6 ABQ82343 | Human NOV |

| | | | | | |
|----|-------|------|-------|------------|-----------|
| 9 | 416 | 75.8 | 4245 | 7 AAL60066 | Human Pco |
| 10 | 416 | 75.8 | 9058 | 7 ACC72051 | BCU0205A |
| 11 | 416 | 75.8 | 9695 | 7 ACC72052 | BCU0205B |
| 12 | 411 | 74.9 | 12879 | 6 ABK92230 | Prostate |
| 13 | 411 | 74.9 | 13202 | 4 AAK51828 | Human pol |
| 14 | 407 | 74.1 | 9729 | 5 AAS14089 | Human FCT |
| 15 | 407 | 74.1 | 9729 | 5 ADB32028 | Human FCT |
| 16 | 407 | 74.1 | 9826 | 5 AAS14085 | Human FCT |
| 17 | 407 | 74.1 | 9826 | 5 ADB32023 | Human FCT |
| 18 | 341 | 62.1 | 466 | 4 ABT21688 | Breast ca |
| 19 | 248.5 | 45.3 | 10242 | 4 ABL29075 | Drosophil |
| 20 | 248.5 | 45.3 | 17131 | 4 ABL29074 | Drosophil |
| 21 | 234.5 | 42.7 | 1764 | 4 ABL04855 | Drosophil |
| 22 | 234.5 | 42.7 | 4925 | 4 ABL04854 | Drosophil |
| 23 | 73 | 13.3 | 2516 | 5 AAF89891 | Peroxisom |
| 24 | 73 | 13.3 | 3008 | 6 ABK90574 | Human PGC |
| 25 | 73 | 13.3 | 3023 | 3 AAA07580 | Human PGC |
| 26 | 73 | 13.3 | 3023 | 6 ABK90559 | Human PGC |
| 27 | 73 | 13.3 | 3023 | 8 ADA26907 | Human PGC |
| 28 | 73 | 13.3 | 6303 | 7 AAL55176 | PGC-1 mut |
| 29 | 73 | 13.3 | 6303 | 7 AAL55139 | Wild-type |
| 30 | 73 | 13.3 | 6306 | 3 AAC75134 | Human ORF |
| 31 | 71 | 12.9 | 2447 | 6 ABL90799 | Human ova |
| 32 | 71 | 12.9 | 2759 | 6 ABQ54635 | Human ova |
| 33 | 71 | 12.9 | 3653 | 6 ABS78730 | DNA encod |
| 34 | 70.5 | 12.8 | 899 | 6 ABN16991 | Human ORF |
| 35 | 70.5 | 12.8 | 866 | 6 ABK65370 | Arabidops |
| 36 | 70.5 | 12.8 | 866 | 9 ADD30950 | Plant yie |
| 37 | 70.5 | 12.8 | 866 | 9 ADE37148 | Plant yie |
| 38 | 70.5 | 12.8 | 1169 | 3 AAC47819 | Arabidops |
| 39 | 70.5 | 12.8 | 4215 | 7 ABX34555 | Human mdd |
| 40 | 70 | 12.8 | 375 | 4 AAI85693 | Human pol |
| 41 | 70 | 12.8 | 816 | 2 AAT67351 | H. pylori |
| 42 | 70 | 12.8 | 855 | 2 AAT68114 | H. pylori |
| 43 | 70 | 12.8 | 984 | 4 AAS53729 | Helicobac |
| 44 | 70 | 12.8 | 984 | 4 AAS53891 | Helicobac |
| 45 | 70 | 12.8 | 2064 | 4 ABL19009 | Drosophil |

ALIGNMENTS

RESULT 1

ABS52100

ID ABS52100 standard; DNA; 8354 BP.

XX ABS52100;

DT 05-NOV-2002 (first entry)

XX Human TEN-M4-like gene.

Human, NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis; cell signal processing; metabolic pathway modulation; metabolic disorder; obesity; diabetes; infectious disease; neurodegenerative disorder; acne; Alzheimer's disease; Parkinson's disease; immune disorder; cancer; haematopoietic disorder; cirrhosis; pancreatitis; learning defect; memory defect; infertility; congenital heart defect; hair growth; pigmentation disorder; endocrine disorder; respiratory disease; health; gastro-intestinal disease; reproductive; neurological disease; bone marrow transplantation; endocrine disease; allergy; inflammation; nephrological disorder; urinary system disorder; age-related disorder; neuropsychiatric disorder; EGF-related protein; SCUB1; TEN-M4; adipocyte complement-related C1q tumour necrosis factor; out at first; beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter; KW type Ia membrane sushi-containing domain; butyrophilin; KW type Ia membrane sushi domain containing; SNP; gene; ds; single nucleotide polymorphism.

OS Homo sapiens.

XX Key Location/Qualifiers
FT variation replace(117,G)
/*tag= a

```
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX WO200257453-A2.
XX
XX 25-JUL-2002.
XX
XX 19-DEC-2001; 2001WO-US050331.
XX
XX 19-DEC-2000; 2000US-0256704P.
XX
XX 20-DEC-2000; 2000US-0257314P.
XX
XX 02-MAY-2001; 2001US-0286153P.
XX
XX 29-MAY-2001; 2001US-0294075P.
XX
XX 24-JUL-2001; 2001US-0307506P.
XX
XX 10-AUG-2001; 2001US-0311590P.
XX
XX 10-AUG-2001; 2001US-0311613P.
XX
XX 29-AUG-2001; 2001US-0315617P.
XX
XX 14-SEP-2001; 2001US-0322358P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Gangolli RA, Patturajan M, Vernet CM, Malyankar UM, Kekuda R;
XX Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
XX Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX WPI; 2002-590744/63.
XX
XX P-PSDB; ABG70388.
XX
XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
XX atherosclerosis, metabolic disorders, diabetes, obesity, infectious
XX disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
XX cancer.
XX
XX Claim 8; Page 50-52; 318pp; English.
XX
XX The present invention relates to new NOVX polypeptides. The invention is
XX useful for treating or preventing a NOVX-associated disorder such as
XX cardiomyopathy or atherosclerosis, where the disorder is related to cell
XX signal processing and metabolic pathway modulation in a subject,
XX preferably human. The invention is also useful for treating metabolic
XX disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
XX disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
XX disorders, haematopoietic disorders and various cancers. The molecules of
XX the invention are also useful for treating or preventing cirrhosis,
XX pancreatitis, learning and memory defects, infertility, congenital heart
XX defects, acne, hair growth, pigmentation disorders, endocrine disorders,
XX respiratory disease, gastro-intestinal diseases, reproductive, health,
XX neurological diseases, bone marrow transplantation, endocrine diseases,
XX allergy and inflammation, nephrological disorders, urinary system
XX disorders, neuropsychiatric disorders and age-related disorders. The
XX present nucleic acid sequence represents a NOVX gene. This sequence
XX encodes a NOVX protein of the invention
XX
XX SQ Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 5,1e-66 Length: 8354
XX Score: 549.00 Matches: 101
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-10-029-020-14_COPY_1100_1200 (1-101) x ABS52100 (1-8354)
XX
XX Qy 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
XX
XX Db 3332 CGCCTCTTCAGAGTGGTTCGTCGAGCCCGCCAGACCTGCTATTATTCATTGGGAC 3391
```

```
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3392 AAGACAGAGCTCTACACAGAGGTGTTTGGGCTTTTCAGAGCCTTTGTTCCGFGGT 3451
QY 41 TyrGluTyrGluSerCysProAspLeuLeuLeuTyrGluLysArgThrThrValLeuGln 60
Db 3452 TATGATATGATCTCTGCCAGATCTAATCTCTGTGGAAAAAGAACACACATGCTGCAG 3511
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 3512 GCGTATGAATTCAGCGCTCCAAAGCTTGAGGATGGAGCTAGACAAACATCATGCGCTC 3571
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3572 AACATTCAAAGTGGTATCTGCACAAAGGAAATGGGAGACCATGTTGTGTCTCAGCAG 3631
QY 101 Pro 101
Db 3632 CCT 3634
RESULT 2
ABS78652
ID ABS78652 standard; cDNA; 8645 BP.
XX
XX AC ABS78652;
XX
XX DT 16-DEC-2002 (first entry)
XX
XX DE Human cDNA encoding CGDD10, INCYTE 7488573CB1.
XX
XX KW Human; ss; Gene; cell growth; differentiation; death; CGDD; cancer;
XX cell proliferative disorder; arteriosclerosis; atherosclerosis;
XX cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
XX polycythaemia vera; primary thrombocytopaenia; developmental disorder;
XX renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
XX neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
XX reproductive disorder; infertility; autoimmune disorder; gout; allergy;
XX inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
XX autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
XX diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
XX multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
XX rheumatoid arthritis.
XX
XX OS Homo sapiens.
XX
XX FN WO200272830-A2.
XX
XX PD 19-SEP-2002.
XX
XX 08-FEB-2002; 2002WO-US003715.
XX
XX 09-FEB-2001; 2001US-0268111P.
XX
XX 23-FEB-2001; 2001US-0271175P.
XX
XX 08-MAR-2001; 2001US-0274503P.
XX
XX 09-MAR-2001; 2001US-0274552P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
XX Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
XX Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
XX
XX WPI; 2002-723356/78.
XX
XX P-PSDB; ABG97359.
XX
XX New human proteins associated with cell growth, differentiation and
XX death, useful for diagnosing, treating or preventing autoimmune or
XX inflammatory disorders (e.g. AIDS, allergy or anaemia), cancer,
XX atherosclerosis or hepatitis.
XX
XX Claim 5; Page 175-178; 181pp; English.
XX
```

CC The invention relates to an isolated polypeptide comprising CGDD1-12
 CC (cell growth, differentiation and death), a naturally occurring amino
 CC acid sequence at least 90% identical to CGDD, a biologically active
 CC fragment or an immunogenic fragment. Also included are the
 CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
 CC comprising a promoter sequence operably linked to the CGDD
 CC polynucleotides, a cell transformed with the recombinant polynucleotide,
 CC a transgenic organism comprising the recombinant polynucleotide, an anti-
 CC CGDD antibody, screening for compounds which bind to/modulate or are
 CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
 CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
 CC agonists and antagonists are useful for diagnosing, treating or
 CC preventing disorders associated with aberrant expression of CGDD,
 CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
 CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
 CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
 CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
 CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
 CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
 CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
 CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
 CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
 CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
 CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
 CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
 CC infections. They are also useful in the assessment of the effects of
 CC exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of proteins associated with CGDD. The present sequence encodes
 CC a CGDD protein
 XX SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 5,35e-66 Length: 8645
 Score: 549.00 Matches: 101
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x ABS78652 (1-8645)

QY 1 ArgLeuPheArgGlySerTyrPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
 Db 3402 CGGCTCTTCAGGAAGTGGTTCGTCGACGCCAGACCTGCTTAATTAATTCATTTGGGAC 3461
 QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
 Db 3462 AAGACAGCGCTACACACAGAGGTGTTGGGCTTTTCAGAGCCTTTGTTTCGGTGGT 3521
 QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGln 60
 Db 3522 TATGAATATGAATCTCCAGATCTAATCTGTGGGAAAAAGACAAACAGTGTCTCAG 3581
 QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyClyTrpSerLeuAspLysHisAlaLeu 80
 Db 3582 GGCTATGAATTTGATGCTCAAGCTTGGAGGATGGAGCCCTAGACAAACATCATGCCCTC 3641
 QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
 Db 3642 AACATTCAAAGTGCATCTCTGCACAAAGGAATGGGAGAACCACTTTGTGTCTCAGCAG 3701
 QY 101 Pro 101
 Db 3702 CCT 3704

RESULT 3
 ABN85378
 ID ABN85378 standard; DNA; 8438 BP.
 XX AC
 XX ABN85378;
 XX 21-OCT-2002 (first entry)

XX DE Human NOV1, TEN-M4 like protein, coding sequence.
 XX KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
 KW Antiallergic; Hemostatic; Anti-HIV; Antidiabetic; Anorectic;
 KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
 KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
 KW Gene therapy; NOV; cancer; heart disease; inflammation;
 KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
 KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
 KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
 KW TEN-M4 like protein; chromosome 11; gene; ds.
 XX OS Homo sapiens.

Key Location/Qualifiers
 CDS 4..8395
 FT /tag= a
 FT /trans except= (pos: 1138..1147,aa:Met)
 FT /product= "NOV1 protein"

WO200255704-A2.

18-JUL-2002.

09-JAN-2002; 2002WO-US000554.

09-JAN-2001; 2001US-0260417P.

10-FEB-2001; 2001US-0260831P.

28-FEB-2001; 2001US-0272338P.

09-MAR-2001; 2001US-0274876P.

18-APR-2001; 2001US-0284704P.

(CURA-) CURAGEN CORP.

Padigar M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;

Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;

Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;

Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;

Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;

Gunther E, Stone DJ;

WPI: 2002-590674/63.

P-PSDB; ABB98401.

NOVX polypeptides and encoding polynucleotides, useful for preventing or
 treating NOVX-associated disorders e.g. cancer, inflammation, or
 Alzheimer's disease, and in chromosome mapping, tissue typing or
 pharmacogenomics.

Claim 9; Page 8-9; 358pp; English.

The present sequence is a coding sequence for a NOV protein. The NOV
 proteins and coding sequences are useful for treating or preventing NOV-
 associated disorders or in the manufacture of a medicament for treating
 the disorders, such as cancer, heart disease, inflammation, autoimmune
 disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
 IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
 (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
 and other wasting disorders associated with chronic diseases. NOV1 is a
 TEN-M4 like protein and the NOV1 gene is localised to chromosome 11

Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.57e-64 Length: 8438
 Score: 538.50 Matches: 101
 Percent Similarity: 99.02% Conservativity: 0
 Best Local Similarity: 99.02% Mismatches: 0
 Query Match: 98.09% Indels: 1
 DB: 6 Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x ABN85378 (1-8438)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerLysTyrPheIleTrpAsp 20
|||
Db 3311 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCAGACCTGCTCTATTATTTTCATTGGGAC 3370

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
|||
Db 3371 AAGACAGACGCTCTACAACAGGAAGTGGTTCAGAGCCCTTTGTTCCGCTGGGT 3430

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
|||
Db 3431 TATGAATATGAATCTGCGCATCTAATCTGCGGAAAAAGAACACACAGTGTGCAG 3490

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
|||
Db 3491 GCGTATGAATTCAGCGCTCCAGCTTGAGGATGGAGCTAGACAAACATATGCCCTC 3550

QY 81 AsnIleGlnSer---GlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGln 99
|||
Db 3551 AACATTCAAGTGGTGGCATCTCGACAAAGGAATGGGAGAACAGTTTGTGTCTCAG 3610

QY 100 GlnPro 101
|||
Db 3611 CAGCCT 3616

RESULT 4

AAC74626

ID AAC74626 standard; cDNA; 453 BP.

XX AC AAC74626;

XX DT 08-FEB-2001 (first entry)

XX DE Human OREF ORF181 polynucleotide sequence SEQ ID NO:361.

XX KW Human; open reading frame; OREF; detection; cytostatic; hepatotropic;

XX KW vulnery; antiparatic; antiparkinsonian; nontropic; neuroprotective;

XX KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;

XX KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

XX KW hypotensive; dermatological; immunosuppressive; antiinflammatory;

XX KW antianemic; antifungal; antirheumatic; antihypertensive;

XX KW neurodegenerative disease; osteoarthritis; graft vs host disease;

XX KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

XX KW cholesterol ester storage; systemic lupus erythematosus; infection;

XX KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;

XX KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;

XX KW bone damage; cartilage damage; antiinflammatory disease; coagulation;

XX KW thrombosis; contraceptive; ss.

XX OS Homo sapiens.

XX PN WO200058473-A2.

XX PD 05-OCT-2000.

XX PF 31-MAR-2000; 2000WO-US008621.

XX PR 31-MAR-1999; 99US-0127607P.

XX PR 02-APR-1999; 99US-0127636P.

XX PR 05-APR-1999; 99US-0127728P.

XX PR 30-MAR-2000; 2000US-00540763.

XX FA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX XX WPI; 2000-602362/57.

XX DR P-PSDB; AAB40417.

XX XX Novel nucleic acids and peptides derived from open reading frame X,

XX XX useful for treating e.g. cancers, proliferative disorders,

XX XX neurodegenerative disorders and cardiovascular disease.

XX Claim 5; Page 617; 5507pp; English.
XX CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human OREF open reading frames 1 to 3161. The OREF
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiparatic; antiparkinsonian; nontropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antihypertensive; antianemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC OREF-associated disorder. The nucleic acids can be used to express OREF
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX Sequence 453 BP; 115 A; 117 C; 116 G; 104 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 2,2e-65 Length: 453
Score: 532.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.90% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AAC74626 (1-453)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerLysTyrPheIleTrpAsp 20
|||
Db 160 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCAGACCTGCTCTATTATTTTCATTGGGAC 219

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
|||
Db 220 AAGACAGACGCTCTACAACAGGAAGTGGTTCAGAGCCCTTTGTTCCGCTGGGT 279

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
|||
Db 280 TATGAATATGAATCTGCGCATCTAATCTGCGGAAAAAGAACACACAGTGTGCAG 339

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
|||
Db 340 GCGTATGAATTCAGCGCTTCCAGCTTGGAGCTGAGCGCTAGACAAACATCATGCCCTC 399

QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSer 98
|||
Db 400 AACATTCAAGTGGCATCTCGACAAAGGAATGGNGAGAACCCAGTTTGTGTCT 453

RESULT 5

ABQ82345

ID ABQ82345 standard; cDNA; 8473 BP.

XX AC ABQ82345;

XX DT 17-DEC-2002 (first entry)

XX DE Human NOV15c encoding cDNA SEQ ID NO:39.

XX KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nontropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; anti allergic; virucide;
KW antianemic; antibacterial; proteoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;

KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 258..8144
FT /*tag= a
FT /product= "NOV15c"
XX

XX WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

XX 29-DEC-2000; 2000US-0258928P.

XX 02-JAN-2001; 2001US-0259415P.

XX 04-JAN-2001; 2001US-0259785P.

XX 20-FEB-2001; 2001US-0269814P.

XX 09-MAR-2001; 2001US-0279863P.

XX 29-MAR-2001; 2001US-0279832P.

XX 29-MAR-2001; 2001US-0279833P.

XX 13-APR-2001; 2001US-0283889P.

XX 18-APR-2001; 2001US-0284447P.

XX 25-APR-2001; 2001US-0286683P.

XX 29-MAY-2001; 2001US-0294080P.

XX 16-AUG-2001; 2001US-0312915P.

XX 17-AUG-2001; 2001US-0313252P.

XX 17-SEP-2001; 2001US-0322899P.

XX 26-NOV-2001; 2001US-0333350P.

XX (CURA-) CURAGEN CORP.

XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;

XX Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;

XX Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;

XX Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;

XX Gunther E, Smithson G, Millet I, Macdougall JR;

XX WPI; 2002-732706/79.

XX P-PSDB; ABP53588.

XX New NOVX polypeptides and polynucleotides useful for treating NOVX-

XX associated disorders, such as cancers, neurological disorders, disorders

XX of vesicular transport, gastrointestinal disorders, and autoimmune

XX diseases.

XX Claim 8; Page 119-121; 444pp; English.

XX The present invention describes novel human proteins designated NOVX,

XX where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,

XX cytotatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,

XX antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,

CC biological sample. The present sequence encodes human NOV15c, which is

CC located on chromosome 4

XX SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;

XX Alignment Scores:

Pred. No.: 1.83e-52 Length: 8473
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x ABQ82345 (1-8473)

QY 1 ArgLeuPheArgLysTyrPheAlaAlaProAspLeuSerTyrTyrPheIleTrrpAsp 20
Db 3150 AGACTCTTCCAAAAGTGGTTTCTGCTCACCACAACTTGGCCTATATCTTTTCATATGGGAT 3209

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3210 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTTATCTGAAGCTGTGTGTGTCAGTTGA 3269

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGln 60
Db 3270 TATGAGTATGAGTCGTTGTTGGACCTGCTGTGGGAAAGAGGACTGCCATTCCTGCAG 3329

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrrpSerLeuAspLysHisAlaLeu 80
Db 3330 GGCATGATTTGGATGGCTCCACATGGTGGTGGCATAGATAACATCAGTGTCTG 3389

QY 81 AsnIleGlnSerGlyLysLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3390 GATGTACAGAACGGTATACTGTACAAGGGAACGGGGAACCAACGTTTCATCTCCACGAG 3449

QY 101 Pro 101

Db 3450 CCT 3452

RESULT 6

ABQ82346

ID ABQ82346 standard; cDNA; 8487 BP.

XX AC ABQ82346;

XX 17-DEC-2002 (first entry)

XX Human NOV15d encoding cDNA SEQ ID NO:41.

XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
XX cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;
XX antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
XX antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
XX cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
XX stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergic;
XX Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
XX diabetes mellitus; Grave's disease; gastrointestinal disorder; infection;
XX ulcerative colitis; gastric disorder; duodenal disorder; infection;
XX autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
XX rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 259..8140
FT /*tag= a
FT /product= "NOV15d"

XX WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

(CURA-) CURAGEN CORP.

PA
XX
PI Spytke KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53587.
XX
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
XX
PS Claim 8; Page 114-117; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytosolic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15b, which is
CC located on chromosome 4
XX
SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;
XX
Alignment Scores:
Pred. No.: 1-88e-52 Length: 8645
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x ABQ82344 (1-8645)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3322 AGACTCTTCAAAAGTGGTTCCTGCTCCACAAAGCTGGCCTATCTTCATGGGAT 3381
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3382 AAAACAGATGATATATACAGAAAGTCTATGGTCTATCTGAAGCTGTGTGTCAGTTGGA 3441
QY 41 TyrGlyTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGln 60
Db 3442 TATGAGTATGAGTGTGTTGGACCTGACTCTGTGGGAAAGAGGACTGCCATTCCTGAG 3501
QY 61 GlyTyrGluIleLeuAspAlaSerLysLysGlyTyrTrpSerLeuAspLysHisAlaLeu 80
Db 3502 GGCTATGAATGGATGGCTCAACATGGTGGCTGGACATAGATAAACATCAGTGTCTG 3561
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3562 GATGTACAGAACGGTATCTACTGTACAGGGAAACGGGAAACACAGTTTCATCTCCACGAG 3621

QY 101 Pro 101
Db 3622 CCT 3624
RESULT 8
ABQ82343
ID ABQ82343 standard; cDNA; 8675 BP.
XX
AC ABQ82343;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15a encoding cDNA SEQ ID NO:35.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 151..8328
/*tag= a
/product= "NOV15a"
/transl_except= {pos:1249..1251,aa:Ser}
W0200262999-A2.
XX
15-AUG-2002.
XX
31-DEC-2001; 2001WO-US049976.
XX
29-DEC-2000; 2000US-0258928P.
XX
02-JAN-2001; 2001US-0259415P.
XX
04-JAN-2001; 2001US-0259785P.
XX
20-FEB-2001; 2001US-0269814P.
XX
09-MAR-2001; 2001US-0279863P.
XX
29-MAR-2001; 2001US-0279832P.
XX
29-MAR-2001; 2001US-0279833P.
XX
13-APR-2001; 2001US-0283889P.
XX
18-APR-2001; 2001US-0284447P.
XX
25-APR-2001; 2001US-0286683P.
XX
29-MAY-2001; 2001US-0294080P.
XX
16-AUG-2001; 2001US-0312915P.
XX
17-AUG-2001; 2001US-031325P.
XX
17-SEP-2001; 2001US-0322699P.
XX
26-NOV-2001; 2001US-0333350P.
XX
(CURA-) CURAGEN CORP.
XX
PI Spytke KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53586.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX

```
PS Claim 8; Page 110-112; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where x is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytosolic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC cytostatic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC antiabietic, antiinflammatory, antirheumatic, antibacterial, fungicide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC procoagulant and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15a, which is
CC located on chromosome 4
XX
SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;
Alignment Scores:
Pred. No.: 1.89e-52 Length: 8675
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x ABQ82343 (1-8675)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3334 AGACTTCCAAAGTGGTTCTCGCTCACCNACTTGGCCTATCTTCTATCGGAT 3393
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3394 AAAACAGATGCATATATCAGAAAGTCTATGGTCTATCTGAGCTGTGTGTCAGTTGGA 3453
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGln 60
Db 3454 TATGAGTATGAGTCGTGTGTGGACCTGACTCTGTGGGAAAGAGAGCTCCATCTTCGAG 3513
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db 3514 GGTATGAAATGGATGGTCCCAACATGGGTGGCTGGACATTAGATAAATCATCAGTGTCTG 3573
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3574 GATGTACAGACGGTATCTGTACAGAGGAAACGGGAAACCGATCTATCTCCAGCAG 3633
QY 101 Pro 101
Db 3634 CCT 3636
RESULT 9
AAL60066
ID AAL60066 standard; cDNA; 4245 BP.
XX
AC AAL60066;
XX
DT 27-AUG-2003 (first entry)
XX
DE Human Pc099 cDNA.
XX
KW Human; differentially regulated protein; prevention; therapy; vaccine;
```

Qy 82 IleGlnSerGlyIleLeuHisGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
 Db 2158 GTTAAAGTGAATCCTACACAAAGGCACTGGGGAACCAAGTCTCTGACCCAGCAGCCT 2217
 RESULT 10
 ACCT2051
 ID ACC72051 standard; DNA; 9058 BP.
 XX
 AC ACC72051;
 XX
 DT 08-JUL-2003 (first entry)
 XX
 DE BCU0205A gene #SEQ ID 79.
 XX
 KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
 KW drug discovery; clinical medicine; forensic medicine; gene;
 KW chromosome 5q33.3; ds.
 XX
 OS Homo sapiens.
 XX
 EN WO2003029421-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-US031287.
 XX
 PR 03-OCT-2001; 2001US-0326526P.
 PR 14-MAY-2002; 2002US-00144194.
 XX
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX
 PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
 XX
 DR WPI; 2003-381623/36.
 DR P-PSDB; ABR58317.
 XX
 PT New isolated human differentially-regulated breast cancer polynucleotide
 PT and polypeptide, useful for diagnosing, staging, prognosticating,
 PT preventing and/or treating diseases and conditions relating to breast
 PT cancer.
 XX
 PS Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.
 XX
 CC The invention relates to isolated polynucleotides which are
 CC differentially-regulated in breast cancer. The methods and compositions
 CC of the present invention are useful for detecting, diagnosing, staging,
 CC monitoring, prognosticating, preventing and/or treating diseases and
 CC conditions relating to breast cancer, and may be used in gene therapy or
 CC antisense therapy. They can also be used in research, drug discovery,
 CC clinical medicine and forensic medicine. Sequences given in records
 CC ACCT2012-ACC72074 represent polynucleotides of the invention that are
 CC differentially-regulated in breast cancer. NOTE: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.34e-47 Length: 9058
 Score: 416.00 Matches: 72
 Percent Similarity: 88.00% Conservative: 16
 Best Local Similarity: 72.00% Mismatches: 12
 Query Match: 75.77% Indels: 0
 DB: Gaps: 7
 US-10-029-020-14_COPY_1100_1200 (1-101) x ACC72051 (1-9058)
 Qy 2 LeupheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
 Db 2699 CTCCTCCAGAGTCAATCCAGGCTTCTCCCAACCTGGCTACACCTTCATCTGGGACAAG 2758
 Qy 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41

Db 2759 ACATGCGTATGGCCAAAGGCTGTATGGACTCTCAGATGCTGTGTCTGTGCGGTTT 2818
 Qy 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValIleuGlnGly 61
 Db 2819 GAATATGAGACCTGTCCAGTCTAATCTCTGGGAGAAAAGACAGCCCTCTCTCAGGGA 2878
 Qy 62 TyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisHisAlaLeuAsn 81
 Db 2879 TTCGAGCTGGACCCCTCCACCTCGGTGGTGGTCCCTAGACAAACACCATCTCAAT 2938
 Qy 82 IleGlnSerGlyIleLeuHisGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
 Db 2939 GTTAAAGTGAATCCTACACAAAGGCACTGGGGAACCAAGTCTCTGACCCAGCAGCCT 2998
 RESULT 11
 ACCT2052
 ID ACC72052 standard; DNA; 9695 BP.
 XX
 AC ACC72052;
 XX
 DT 08-JUL-2003 (first entry)
 XX
 DE BCU0205B gene #SEQ ID 81.
 XX
 KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
 KW drug discovery; clinical medicine; forensic medicine; gene;
 KW chromosome 5q33.3; ds.
 XX
 OS Homo sapiens.
 XX
 EN WO2003029421-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-US031287.
 XX
 PR 03-OCT-2001; 2001US-0326526P.
 PR 14-MAY-2002; 2002US-00144194.
 XX
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX
 PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
 XX
 DR WPI; 2003-381623/36.
 DR P-PSDB; ABR58318.
 XX
 PT New isolated human differentially-regulated breast cancer polynucleotide
 PT and polypeptide, useful for diagnosing, staging, prognosticating,
 PT preventing and/or treating diseases and conditions relating to breast
 PT cancer.
 XX
 PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.
 XX
 CC The invention relates to isolated polynucleotides which are
 CC differentially-regulated in breast cancer. The methods and compositions
 CC of the present invention are useful for detecting, diagnosing, staging,
 CC monitoring, prognosticating, preventing and/or treating diseases and
 CC conditions relating to breast cancer, and may be used in gene therapy or
 CC antisense therapy. They can also be used in research, drug discovery,
 CC clinical medicine and forensic medicine. Sequences given in records
 CC ACCT2012-ACC72074 represent polynucleotides of the invention that are
 CC differentially-regulated in breast cancer. NOTE: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.67e-47 Length: 9695
 Score: 416.00 Matches: 72
 Percent Similarity: 88.00% Conservative: 16

Best Local Similarity: 72.00% Mismatches: 12
Query Match: 75.7% Indels: 0
DB: 7 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x ACC72052 (1-9695)

QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
3336 CTTCTCCAGAGTCATCCAGGCTTCCCAACCTGGCTACACCTTCATCTGGGACAAG 3395
QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
3396 ACAGATGCGTATGGCCAAAGGGGTATGGACTTCAGATGCTGTGTGTCGTGCGGGTTT 3455
QY 42 GluTyrGluSerCysProAspLeuLeuTyrGluLysArgThrThrValLeuGlnGly 61
3456 GAATATGAGACCTGCTCCCAAGTCTAATCTCTGGGAGAAAGGACACCTCTCTCAGGGA 3515
QY 62 TyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeuAsn 81
3516 TTCGAGCTGAGCCCTCCCAACCTCGTGGTGGTCTAGACAAACACCATCTCTCAAT 3575
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlnGluAsnGlnPheValSerGlnGlnPro 101
3576 GTTAAAGTGAATCCTCACAAAGGCACTGGGGAAACCCAGTCTCTGACCCAGCAGCCT 3635

RESULT 12
ABK92230
ID ABK92230 standard; DNA; 12879 BP.
XX
AC ABK92230;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated DNA sequence #116.

XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
KW gene therapy; gene; ds.
XX Mammalia.
XX WO200230268-A2.
XX 18-APR-2002.
XX 12-OCT-2001; 2001WO-US032045.
XX 13-OCT-2000; 2000US-00687576.
XX 08-DEC-2000; 2000US-00733288.
XX 08-DEC-2000; 2000US-00733742.
XX 24-JAN-2001; 2001US-0263957P.
XX 16-MAR-2001; 2001US-0276791P.
XX 16-MAR-2001; 2001US-0276888P.
XX 06-APR-2001; 2001US-0281922P.
XX 24-APR-2001; 2001US-0286214P.
XX 30-APR-2001; 2001US-00847046.
XX 04-MAY-2001; 2001US-0288589P.
XX (BOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX WPI; 2002-471335/50.
XX p-PSDB; ABG61913.

XX Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
XX Claim 22; Page 394-397; 436pp; English.
XX
XX The present invention relates to methods of detecting a prostate cancer-

CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences
XX
XX Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;

SQ

Alignment Scores:
Pred. No.: 2,77e-46 Length: 12879
Score: 411.00 Matches: 73
Percent Similarity: 85.15% Conservative: 13
Best Local Similarity: 72.28% Mismatches: 15
Query Match: 74.86% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x ABK92230 (1-12879)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
3257 CGACTCACACAGAAGTGGTTCCCGCCCAATTAATCTTGTCTACACATTTGCTTGAAC 3316
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
3317 AGACCCGATATCTATGACAGAGGTTGGGGCTGGCAGAGGCTTTGGTATCTGTGGGA 3376
QY 41 TyrGluTyrGluSerCysProAspLeuLeuTyrGluLysArgThrThrValLeuGln 60
3377 TATGAATATGAAAGTGGCTGACTTATTTCTCTGGGAGCAAGAGCAGTCGTTTTTACAA 3436
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
3437 GGTTTTGAGATGATGCTTCTAACCTAGGAGACTGGTCTTTGTAATAGCATCATCTTTG 3496
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
3497 AATCCTCAAGTGAATCATACATAAAGGGAATGGAGAAATATGTTTCATTTCCAGCAG 3556
QY 101 Pro 101
3557 CCC 3559

RESULT 13

AAK51828
ID AAK51828 standard; cDNA; 13202 BP.
XX
AC AAK51828;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 373.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX Homo sapiens.
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX

PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
DR P-PSDB; AAM78695.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
XX Claim 1; Page 1414-1426; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78333-AAK80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,86e-46 Length: 13202
Score: 411.00 Matches: 73
Percent Similarity: 85.15% Conservative: 13
Best Local Similarity: 72.28% Mismatches: 15
Query Match: 74.86% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AAK51828 (1-13202)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTrpPheIleTrpAsp 20
Db 3583 CGAGCTCACAGAGAAGTGGTTCCGCGCGCAATTAATCTGCTACACATTTGCTGGAAC 3642

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3643 AAGACCGATATCATGGACAGAGAGTTGGGCGCTCGCAGAGGCTTTGGTATCTGTGGGA 3702

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGln 60
Db 3703 TATGAATATGAACGTCCTCGATTTATCTCTGGGAGCAAGAGCAGTCGTTTACAA 3762

QY 61 GlyTyrGluLeuAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 3763 GGTTTTGAGATGGATGCTTCTTAACCTAGGAGAGCTGGTCTTTGAATAGCATCATTTTG 3822

QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlnLysGluAsnGlnPheValSerGlnGln 100
Db 3823 AATCCTCAAGTGAATCATACATAAAGGAATGAGAAATATGTTCATTTCCACAG 3882

QY 101 Pro 101
|||

Db 3883 CCC 3885
RESULT 14
AAS14089
ID AAS14089 standard; DNA; 9729 BP.
XX
XX AAS14089;
XX
XX 18-DEC-2001 (first entry)
XX
XX Human FCTR3f DNA sequence.
XX
XX Human; FCTRX; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
KW astrocytoma; congenital neonatal alloimmune thrombocytopaenia; infection;
KW neurological disorder; neurodegenerative disorders; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greonow type I; Corneal dystrophy-lattice type I;
KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW anti-allergic; antiasthmatic; antiinfertility; antiinflammatory;
KW antidiabetic; protozoicide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3a; neurestin-like protein; FCTR3f.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT 5'UTR 1..209 /*tag= b
FT CDS 210..8384 /*tag= a
FT FT /product= "Human FCTR3f"
FT 3'UTR 8385..9729 /*tag= c
FT
FT
XX WO200166747-A2.
XX
XX 13-SEP-2001.
XX
XX 05-MAR-2001; 2001WO-US007160.
XX
XX 03-MAR-2000; 2000US-0186592P.
XX 03-MAR-2000; 2000US-0186718P.
XX 06-MAR-2000; 2000US-0187293P.
XX 06-MAR-2000; 2000US-0187294P.
XX 17-MAR-2000; 2000US-0190400P.
XX 07-APR-2000; 2000US-0196018P.
XX 03-JAN-2001; 2001US-0259548P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Verna CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
XX WPI; 2001-596837/67.
XX P-PSDB; AAU08681.
XX
XX Novel polypeptides designated as FCTRX polypeptides, useful in detection,
PT prevention and treatment of a broad range of pathological states.
XX
XX Claim 9; Page 37-39; 215pp; English.
XX
XX The invention relates to human FCTRX polypeptides, FCTR1-FCTR7, and the
CC nucleic acids encoding them. These sequences are useful for the treatment
CC or prevention of numerous disorders including myelogenous leukaemia,
CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
CC alloimmune thrombocytopaenia, neurological disorders, neurodegenerative

CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
CC myelodysplastic syndrome, mental health conditions, immunological
CC disorders, allergy and infection, bronchial asthma, Avellino type
CC eosinophilia, lung diseases, reproductive disorders, infertility, male
CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
CC Corneal dystrophy-Groenouw type I, Corneal dystrophy-lattice type I and
CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
CC the FCTR3a homologue FCTR3f
XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.86e-46 Length: 9729
Score: 407.00 Matches: 71
Percent Similarity: 87.00% Conservative: 16
Best Local Similarity: 71.00% Mismatches: 13
Query Match: 74.13% Indels: 0
DB: 5 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AAS14089 (1-9729)

QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyTyPheIleTrpAspLys 21
DB 3384 CTCCTCCAGAGTCATTCAGGCTTCTCCACCTGGCTCCACCTTCATCTGGGACAAG 3443
QY 22 ThrAspValTyAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
DB 3444 ACAGATGCGTATGGCCAAAGGGTGTATGGACTCTCAGACTCTCAGTGTGTGTCTGCGGGTTT 3503
QY 42 GluTyrGluSerCysProAspLeuLeuTyrGluLysArgThrThrValLeuGlnGly 61
DB 3504 GAATATGAGACCTGCTCCCTAGTCTTAATCTCTGGGAGAAAGACAGCCCTCTCTCAGGGA 3563
QY 62 TyrGluIleAspAlaSerLysLeuGlyGlyTyrPheSerLeuAspLysHisAlaLeuAsn 81
DB 3564 TTCGAGCTGGACCCCTCCACCTCGTGGTGGTCCCTAGACACACCATCTCAT 3623
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
DB 3624 GTTAAAGTGAATCCTACACAAAGGCACTGGGGAAACCAAGTCTCTGACCCAGCAGCCT 3683

RESULT 15

ADB32028

ID ADB32028 standard; cDNA; 9729 BP.

XX AC

XX DT

XX 04-DEC-2003 (first entry)

XX Human FCTR3f cDNA.

XX DE

KW Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW lattice corneal dystrophy.

XX Homo sapiens.

OS US2003087816-A1.

FN

XX 08-MAY-2003.
PD
XX 05-MAR-2001; 2001US-00800198.
XX
XX 03-MAR-2000; 2000US-0186592P.
PA (VERM/) VERMET C.
PA (PERN/) FERNANDES E.
PA (SHIM/) SHIMKETS R.
PA (HERR/) HERRMANN J.
PA (MAJU/) MAJUMDER K.
PA (MACD/) MACDOUGALL J.
PA (MISH/) MISHRA V.
PA (MEZE/) MEZES P S.
PA (RAST/) RASTELLI L.
XX Vernet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
PI MacDougall J, Mishra V, Mezes PS, Rastelli L;
XX WPI: 2003-625633/59.
DR P-PSDB; ADB32029.
XX
PT New FCTR3 polypeptide and encoding polynucleotide, useful for preventing
or treating FCTR3-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
diseases.
XX
PS Claim 9; Page 34-37; 155pp; English.
XX The invention relates to FCTR3 polypeptides and the polynucleotides
encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTR3 polypeptide, such as colorectal cancer.
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, renal cell
CC tumours, mammary tumours, human gliomas, astrocytomas, clear cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, nerve trauma, familial myelodysplastic
CC neurodegenerative disorders, Gardner syndrome, mental health
CC syndrome, Charcot-Marie-Tooth neuropathy, allergy and infection, asthma, lung
CC conditions, immunological disorders, reproductive disorders, deafness, glycoprotein
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTR3 polypeptide of the invention.
XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.86e-46 Length: 9729
Score: 407.00 Matches: 71
Percent Similarity: 87.00% Conservative: 16
Best Local Similarity: 71.00% Mismatches: 13
Query Match: 74.13% Indels: 0
DB: 5 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x ADB32028 (1-9729)

QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyTyPheIleTrpAspLys 21
DB 3384 CTCCTCCAGAGTCATTCAGGCTTCTCCACCTGGCTCCACCTTCATCTGGGACAAG 3443
QY 22 ThrAspValTyAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
DB 3444 ACAGATGCGTATGGCCAAAGGGTGTATGGACTCTCAGACTCTCAGTGTGTGTCTGCGGGTTT 3503
QY 42 GluTyrGluSerCysProAspLeuLeuTyrGluLysArgThrThrValLeuGlnGly 61
DB 3504 GAATATGAGACCTGCTCCCTAGTCTTAATCTCTGGGAGAAAGACAGCCCTCTCTCAGGGA 3563

```
QY      62  TyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisHisAlaLeuAsn 81
Db      3564  TTCAGCTGGACCCCTCCAACTCGGTGGCTGGTCCCTAGACAAACACCATCCTCAAT 3623

QY      82  IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
Db      3624  GTTAAAGTGGAAATCCTACACAAAGGCACTGGGGAAACCCAGTTCTTGACCCAGCAGCCT 3683
```

Search completed: August 14, 2004, 02:29:50
Job time : 197.323 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 202.425 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_750_850
Perfect score: 617
Sequence: 1 WMGACDQQRACHPRCAEHGT.....DTSMETACGDSKNDGDGLV 101

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SURFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCI=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10029020@cgn2_1.1.2156@runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=130 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
|------------|-------|-------------|--------|-------|-------------|

ALIGNMENTS

RESULT 1
US-10-383-201-49
; Sequence 49, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

| | | | | | | |
|----|-------|-------|--------|----|---------------------|---------------------|
| 1 | 617 | 100.0 | 1399 | 13 | US-10-383-201-49 | Sequence 49, Appl |
| 2 | 617 | 100.0 | 1399 | 13 | US-10-383-201-59 | Sequence 59, Appl |
| 3 | 617 | 100.0 | 8354 | 13 | US-10-383-201-43 | Sequence 43, Appl |
| 4 | 617 | 100.0 | 8354 | 13 | US-10-029-020-13 | Sequence 13, Appl |
| 5 | 595 | 96.4 | 8355 | 13 | US-10-383-201-55 | Sequence 55, Appl |
| 6 | 573 | 93.8 | 8438 | 13 | US-10-042-865-1 | Sequence 1, Appl |
| 7 | 555.5 | 90.0 | 1371 | 13 | US-10-383-201-51 | Sequence 51, Appl |
| 8 | 515 | 83.5 | 8473 | 17 | US-10-038-854-39 | Sequence 39, Appl |
| 9 | 515 | 83.5 | 8645 | 17 | US-10-038-854-37 | Sequence 37, Appl |
| 10 | 515 | 83.5 | 8675 | 17 | US-10-038-854-35 | Sequence 35, Appl |
| 11 | 509 | 82.5 | 829 | 13 | US-10-383-201-47 | Sequence 47, Appl |
| 12 | 509 | 82.5 | 829 | 13 | US-10-383-201-61 | Sequence 61, Appl |
| 13 | 509 | 82.5 | 1392 | 13 | US-10-383-201-45 | Sequence 45, Appl |
| 14 | 509 | 82.5 | 1392 | 13 | US-10-383-201-53 | Sequence 53, Appl |
| 15 | 509 | 82.5 | 1476 | 13 | US-10-383-201-41 | Sequence 41, Appl |
| 16 | 478.5 | 77.6 | 8487 | 17 | US-10-038-854-41 | Sequence 41, Appl |
| 17 | 464.5 | 75.3 | 12880 | 16 | US-10-295-027-927 | Sequence 927, Appl |
| 18 | 454 | 73.6 | 9826 | 9 | US-09-808-602-7 | Sequence 7, Appl |
| 19 | 454 | 73.6 | 9826 | 10 | US-09-800-198-7 | Sequence 7, Appl |
| 20 | 452 | 73.3 | 8409 | 9 | US-09-808-602-79 | Sequence 79, Appl |
| 21 | 452 | 73.3 | 8409 | 10 | US-09-800-198-67 | Sequence 67, Appl |
| 22 | 448.5 | 72.7 | 8689 | 9 | US-09-808-602-78 | Sequence 78, Appl |
| 23 | 448.5 | 72.7 | 8689 | 10 | US-09-800-198-66 | Sequence 66, Appl |
| 24 | 448.5 | 72.7 | 8797 | 9 | US-09-808-602-74 | Sequence 74, Appl |
| 25 | 448.5 | 72.7 | 8797 | 9 | US-09-808-602-77 | Sequence 77, Appl |
| 26 | 448.5 | 72.7 | 8797 | 10 | US-09-800-198-62 | Sequence 62, Appl |
| 27 | 448.5 | 72.7 | 8797 | 10 | US-09-800-198-65 | Sequence 65, Appl |
| 28 | 448.5 | 72.7 | 9729 | 9 | US-09-808-602-12 | Sequence 12, Appl |
| 29 | 448.5 | 72.7 | 9729 | 10 | US-09-800-198-12 | Sequence 12, Appl |
| 30 | 447.5 | 72.5 | 802 | 13 | US-10-383-201-57 | Sequence 57, Appl |
| 31 | 447.5 | 72.5 | 8575 | 13 | US-10-072-012-143 | Sequence 143, Appl |
| 32 | 430.5 | 69.8 | 447 | 9 | US-09-833-381-1072 | Sequence 1072, Appl |
| 33 | 330.5 | 53.6 | 9058 | 16 | US-10-144-194A-79 | Sequence 79, Appl |
| 34 | 330.5 | 53.6 | 9695 | 16 | US-10-144-194A-81 | Sequence 81, Appl |
| 35 | 216 | 35.0 | 2496 | 9 | US-09-808-602-75 | Sequence 75, Appl |
| 36 | 216 | 35.0 | 2496 | 10 | US-09-800-198-63 | Sequence 63, Appl |
| 37 | 199 | 32.3 | 2860 | 15 | US-10-029-386-24765 | Sequence 24765, A |
| 38 | 199 | 32.3 | 600 | 15 | US-10-029-386-11065 | Sequence 11065, A |
| 39 | 169 | 27.4 | 13268 | 15 | US-10-007-926A-47 | Sequence 47, Appl |
| 40 | 169 | 27.4 | 13857 | 15 | US-10-037-270-75 | Sequence 75, Appl |
| 41 | 169 | 27.4 | 13857 | 16 | US-10-117-722-75 | Sequence 75, Appl |
| 42 | 169 | 27.4 | 100267 | 13 | US-10-240-425-1470 | Sequence 1470, Appl |
| 43 | 161 | 26.1 | 536 | 15 | US-10-029-386-10449 | Sequence 10449, A |
| 44 | 161 | 26.1 | 1251 | 15 | US-10-029-386-24152 | Sequence 24152, A |
| 45 | 156 | 25.3 | 7560 | 10 | US-09-921-406C-17 | Sequence 17, Appl |

;
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 49
; LENGTH: 1399
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1399)
US-10-383-201-49

Alignment Scores:
Pred. No.: 9.8e-64 Length: 1399
Score: 617.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-49 (1-1399)

QY 1 TrpMetGlyAlaAalaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||
Db 578 TGGATGGGGGCGAGCTGCGACCGGGCGCTGCCACCGCGCTGTGCCAGCATGGGACC 637
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|||
Db 638 TGCCCGCAGCGCAAGTGGCAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGCT 697
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
|||
Db 698 CACTATCTGGATAGGTAGTTAAAGAGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 757
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
|||
Db 758 TGTACCTTAGACCTGAATGGTGGCAGCTCGCTGCTGCCAGCTGGCTGGAGAGGAGCTGC 817
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
|||
Db 818 TGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCTG 877
QY 101 Val 101
|||
Db 878 GTG 880

RESULT 2

US-10-383-201-59
; Sequence 59, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755

;
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 59
; LENGTH: 1399
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(1399)
US-10-383-201-59

Alignment Scores:
Pred. No.: 9.8e-64 Length: 1399
Score: 617.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-59 (1-1399)

QY 1 TrpMetGlyAlaAalaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||
Db 578 TGGATGGGGGCGAGCTGCGACCGGGCGCTGCCACCGCGCTGTGCCAGCATGGGACC 637
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|||
Db 638 TGCCCGCAGCGCAAGTGGCAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGCT 697
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
|||
Db 698 CACTATCTGGATAGGTAGTTAAAGAGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 757
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
|||
Db 758 TGTACCTTAGACCTGAATGGTGGCAGCTCGCTGCTGCCAGCTGGCTGGAGAGGAGCTGC 817
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
|||
Db 818 TGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCTG 877
QY 101 Val 101
|||
Db 878 GTG 880

RESULT 3

US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779

; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 7,93e-63 Length: 8354
Score: 617.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-43 (1-8354)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2282 TGGATGGGGCAGCTGCGACGCGGCGCTGCACCCGGCGCTGCGCAGCATGGGACC 2341
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2342 TGGCGGACGGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCAACTGCACCATCGCT 2401
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2402 CACTATCTGATAGGTAGTTAAAGAGGTGGCTGGTGTGCAATGGCAACGGCAGA 2461
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2462 TGTACCTTAGACCTGAATGCTGGCTGCTGCCAGCTGGGCTGGAGAGAGCTGGC 2521
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2522 TGTGACACTTCCATGGAGACTGCTCGGCTGGAGCAAGCAACAATGATGGAGATGGCCTG 2581
QY 101 Val 101
Db 2582 GTG 2584

RESULT 4

US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gargolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613

; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-029-020-13

Alignment Scores:
Pred. No.: 7,93e-63 Length: 8354
Score: 617.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-029-020-13 (1-8354)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2282 TGGATGGGGCAGCTGCGACGCGGCGCTGCACCCGGCGCTGCGCAGCATGGGACC 2341
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2342 TGGCGGACGGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCAACTGCACCATCGCT 2401
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2402 CACTATCTGATAGGTAGTTAAAGAGGTGGCTGGTGTGCAATGGCAACGGCAGA 2461
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2462 TGTACCTTAGACCTGAATGCTGGCTGCTGCCAGCTGGGCTGGAGAGAGCTGGC 2521
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2522 TGTGACACTTCCATGGAGACTGCTCGGCTGGAGCAAGCAACAATGATGGAGATGGCCTG 2581
QY 101 Val 101
Db 2582 GTG 2584

RESULT 5

US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26

```
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:
Pred. No.: 3,298-60 Length: 8355
Score: 595.00 Matches: 101
Percent Similarity: 96.19% Conservative: 0
Best Local Similarity: 96.19% Mismatches: 0
Query Match: 96.43% Indels: 4
DB: 13 Gaps: 2

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-55 (1-8355)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
DB 2239 TGGATGGGGGAGCTTGGACGAGGGGGCTGCCACCCCGCTGTCGCGAGCATGGGACC 2298
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle--- 39
DB 2299 TGCCGCGACGCAAGTGGAGTGCAGCCCTGGCTGGATGGCGAAGTGCACCTGCACCATCTCC 2358

QY 40 ---AlaHisTyrLeuAspArgValVallys-----GluGlyCysProGlyLeuCysAsn 56
DB 2359 CTAGTCTCACTATCTGGATAGGTAGTTAACTTTTCAGAGGGTGGCCCTGGTGTGCAAT 2418
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
DB 2419 GGCACGCGCAGATGTACCTTAGACCTGATGTTGGCACTGGCTGTGCCAGCTGGGCTGG 2478
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerIysAspAsnAsp 96
DB 2479 AGAGGAGCTGGTGTGACACTTCCATGGAGACTGCTGGGTGACAGCAAGACAATGAT 2538
QY 97 GlyAspGlyLeuVal 101
DB 2539 GGAGATGGCCTGGTG 2553

RESULT 6
US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T

; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Refence L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1

Alignment Scores:
Pred. No.: 2,678-58 Length: 8438
Score: 579.00 Matches: 98
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 93.33% Mismatches: 2
Query Match: 93.84% Indels: 4
DB: 13 Gaps: 2

US-10-029-020-14_COPY_750_850 (1-101) x US-10-042-865-1 (1-8438)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
DB 2249 TGGACGGGCGCCAGCTGTAATCAGAGAGCTGCCACCCCGCTGTCGCGAGCATGGGACC 2308
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle--- 39
DB 2309 TGCCGCGACGCGCAAGTGGAGTGCAGCCCTGGCTGGATGGCGAAGTGCACCATCTCC 2368
QY 40 ---AlaHisTyrLeuAspArgValVallys-----GluGlyCysProGlyLeuCysAsn 56
DB 2369 CTAGTCTCACTATCTGGATAGGTAGTTAACTTTTCAGAGGGTGGCCCTGGTGTGCAAT 2428
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
DB 2429 GGCACGCGCAGATGTACCTTAGACCTGAATGGTTGGCACTGGCTGTGCCAGCTGGGCTGG 2488
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerIysAspAsnAsp 96
DB 2489 AGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCTGGGTGACAGCAAGACAATGAT 2548
QY 97 GlyAspGlyLeuVal 101
DB 2549 GGAGATGGCCTGGTG 2563
```

RESULT 7

US-10-383-201-51

Sequence 51, Application US/10383201

Publication No. US20040029226A1

GENERAL INFORMATION:

APPLICANT: Alsebrook II, John et al.

TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS

FILE REFERENCE: 21402-568A

CURRENT APPLICATION NUMBER: US/10/383,201

CURRENT FILING DATE: 2003-03-06

PRIOR APPLICATION NUMBER: 10/029020

PRIOR FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: 60/365,984

PRIOR FILING DATE: 2002-03-20

PRIOR APPLICATION NUMBER: 60/372,022

PRIOR FILING DATE: 2002-04-12

PRIOR APPLICATION NUMBER: 60/389,143

PRIOR FILING DATE: 2002-06-14

PRIOR APPLICATION NUMBER: 60/391,779

PRIOR FILING DATE: 2002-06-26

PRIOR APPLICATION NUMBER: 60/410,755

PRIOR FILING DATE: 2002-09-13

PRIOR APPLICATION NUMBER: 60/412,957

PRIOR FILING DATE: 2002-09-23

PRIOR APPLICATION NUMBER: 10/051,874

PRIOR FILING DATE: 2002-01-16

PRIOR APPLICATION NUMBER: 60/366,928

PRIOR FILING DATE: 2002-03-22

PRIOR APPLICATION NUMBER: 10/055,877

PRIOR FILING DATE: 2002-01-22

NUMBER OF SEQ ID NOS: 155

SOFTWARE: CuraseqList version 0.1

SEQ ID NO 51

LENGTH: 1371

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (2)..(1369)

US-10-383-201-51

Alignment Scores:

Pred. No.: 2e-56

Score: 555.50

Percent Similarity: 91.09%

Best Local Similarity: 91.09%

Query Match: 90.03%

DB: 13

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-51 (1-1371)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20

Db 578 TGGATGGGGGACGCTCGACACCGGGCTCGACCCGGCGCTGTGCCGAGCATGGGACC 637

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40

Db 638 TGGCGGACGGCAAGTCGAGTCGAGCCCTGGCTGGGAATGGCAACACTGCACCATC--- 694

QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60

Db 695 -----GAGGGTTGCCCTGGGTGTGCAATGGCAATGGCAGA 730

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80

Db 731 TGTACCTTAGACTGAATGTTGGCACTGCGCTGTGCCAGCTGGGCTGGAGAGAGCTGGC 790

QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerIysAspAsnAspGlyAspGlyLeu 100

Db 791 TGTGACACTTCCATGGAGACTGCGCTGGGTGACAGCAAGACAAATGATGGAGATGSCCTG 850

QY 101 Val 101

|||

```
DB: 17 Gaps: 1
US-10-029-020-14_COPY_750_850 (1-101) x US-10-038-854-39 (1-8473)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2082 TGGACGGGGCCAGCCTGTAATCAGAGAGCGCTGCCACCCCGCTGTGCCAGCAGCGGACC 2141
QY 21 CysArgAspGlyVysCysGluCysSerProGlyTTPAsnGlyGluHisCysThrIleAla 40
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2142 TGCACAGGATGGCAATGGAATGTAATGCAGCCAGCGGCTGGAATGGAGAGCGTATCGCT 2201
QY 41 HisTyrLeuAspArgValVal-----LysGluGlyCysProGlyLeu 54
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2202 CACTATTGGATAAGATAGTTAAAGACACAGATAGGATATAAGAGAGGGTGTCTGTCTG 2261
QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTTPHisCysValCysGlnLeu 74
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2262 TGCACACAGCAATGGAAGATGTACCTCGGACCAAAATGGCGGACATTTGTGTGCCAGCCT 2321
QY 75 GlyTTPArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp 94
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2322 CGATGGAGAGGAGCGAGGCTGTACCGTAGCCATGGAGACTCTTTTGACAGATAGCAAGGAC 2381
QY 95 AsnAspGlyAspGlyLeuVal 101
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2382 AATGAGGGGATGGACTCATT 2402

RESULT 9
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832

; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 8645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-37

Alignment Scores:
Pred. No.: 1,148-50 Length: 8645
Score: 515.00 Matches: 84
Percent Similarity: 85.05% Conservative: 7
Best Local Similarity: 78.50% Mismatches: 10
Query Match: 83.47% Indels: 6
DB: 17 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x US-10-038-854-37 (1-8645)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2254 TGGACGGGGCCAGCCTGTAATCAGAGAGCGCTGCCACCCCGCTGTGCCAGCAGCGGACC 2313
QY 21 CysArgAspGlyVysCysGluCysSerProGlyTTPAsnGlyGluHisCysThrIleAla 40
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2314 TGCACAGGATGGCAATGGAATGTAATGCAGCCAGCGGCTGGAATGGAGAGACTGCACTATCGCT 2373
QY 41 HisTyrLeuAspArgValVal-----LysGluGlyCysProGlyLeu 54
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2374 CACTATTGGATAAGATAGTTAAAGACACAGATAGGATATAAGAGAGGGTGTCTGTCTG 2433
QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTTPHisCysValCysGlnLeu 74
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2434 TGCACACAGCAATGGAAGATGTACCTCGGACCAAAATGGCGGACATTTGTGTGCCAGCCT 2493
QY 75 GlyTTPArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp 94
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2494 GGATGGAGAGGAGAGCGGCTGTGACGTAGCCATGGAGACTCTTTTGACAGATAGCAAGGAC 2553
QY 95 AsnAspGlyAspGlyLeuVal 101
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2554 AATGAGGGGATGGACTCATT 2574

RESULT 10
US-10-038-854-35
; Sequence 35, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
```

APPLICANT: Gangolli, Esha A
APPLICANT: Guo, Xiaojia S
APPLICANT: Shenoy, Suresh G
APPLICANT: Rastelli, Luca
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc
APPLICANT: Burgess, Catherine E
APPLICANT: Edinger, Shlomit R
APPLICANT: Ellerman, Karen
APPLICANT: Gunther, Erik
APPLICANT: Smithson, Glenda
APPLICANT: Millet, Isabelle
APPLICANT: MacDougall, John R
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-230
CURRENT APPLICATION NUMBER: US/10/038,854
CURRENT FILING DATE: 2003-01-22
PRIOR APPLICATION NUMBER: 60/258,928
PRIOR FILING DATE: 2000-12-29
PRIOR APPLICATION NUMBER: 60/259,415
PRIOR FILING DATE: 2001-01-02
PRIOR APPLICATION NUMBER: 60/259,785
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: 60/269,814
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/279,832
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,833
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/279,863
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/283,889
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,447
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/286,683
PRIOR FILING DATE: 2001-04-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 411
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 35
LENGTH: 8675
TYPE: DNA
ORGANISM: Homo sapiens
US-10-038-854-35

Alignment Scores:
Pred. No.: 1.14e-50 Length: 8675
Score: 515.00 Matches: 84
Percent Similarity: 85.05% Conservative: 7
Best Local Similarity: 78.50% Mismatches: 10
Query Match: 83.47% Indels: 6
DB: 17 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x US-10-038-854-35 (1-8675)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2266 TGACGGGCCCGCCCTGTATACAGAGCGCTGCCACCCCGCTGTGCCGAGCAGCGGACC 2325
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2326 TGCAAGGATGGCAAGTGTGAATCAGCCAGCGGTGTGAATGAGAGGACTGCATATCGCT 2385
QY 41 HisTyLeuAspArgValVal-----LysGluGlyCysProGlyLeu 54
Db 2386 CACTATTGGATAAGATAGTTAAAGACACATAGCATATAAAGAGGTTGTCTGCTG 2445
QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
Db 2446 TGCAACAGCAGTGAAGATGTACCCCTGGACCAAAATGGCGGACATTGTGTGCCAGCCT 2505
QY 75 GlyTrpArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp 94

Db 2506 GGATGGAGAGGAGCAGCTGTGACGTAGCCATGGAGACTCTTTTCACAGATAGCAGGAC 2565
QY 95 AsnAspGlyAspGlyLeuVal 101
Db 2566 AATGAAGGGGATGACTCATT 2586
RESULT 11
US-10-383-201-47
Sequence 47, Application US/10383201
Publication No. US20040029226A1
GENERAL INFORMATION:
APPLICANT: Alsbrook II, John et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
FILE REFERENCE: 21402-568A
CURRENT APPLICATION NUMBER: US/10/383,201
CURRENT FILING DATE: 2003-03-06
PRIOR APPLICATION NUMBER: 10/029020
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/365,984
PRIOR FILING DATE: 2002-03-20
PRIOR APPLICATION NUMBER: 60/372,022
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/389,143
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 60/391,779
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: 60/410,755
PRIOR FILING DATE: 2002-09-13
PRIOR APPLICATION NUMBER: 60/412,957
PRIOR FILING DATE: 2002-09-23
PRIOR APPLICATION NUMBER: 10/051,874
PRIOR FILING DATE: 2002-01-16
PRIOR APPLICATION NUMBER: 60/366,928
PRIOR FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 10/055,877
PRIOR FILING DATE: 2002-01-22
NUMBER OF SEQ ID NOS: 155
SOFTWARE: CuraseqList version 0.1
SEQ ID NO 47
LENGTH: 829
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(829)
US-10-383-201-47

Alignment Scores:
Pred. No.: 3.8e-51 Length: 829
Score: 509.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 82.50% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-47 (1-829)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 578 TGATGGGGGCGACCTGCCAGCCAGCGGGCTGCCACCCGCGCTGTGCCGAGCATGGACC 637
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 638 TGCCGCGACGCGAAGTGCAGTGCAGCCCTGTGGATGGCGAACACTGCACATCGCT 697
QY 41 HisTyLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 698 CACTATCTCGATAGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 757
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 758 TGTACCTTAGACTGTAATGTTGGCACTGCGCTGCCAGCTGGGCTGGAGAGGAGCTGGC 817

```
QY      81 Cys 81
Db      818 TGT 820

RESULT 12
US-10-383-201-61
; Sequence 61, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 61
; LENGTH: 829
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(820)
US-10-383-201-61

Alignment Scores:
Pred. No.:      3.8e-51      Length:      829
Score:          509.00      Matches:      81
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    82.50%      Indels:      0
DB:             13          Gaps:          0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-61 (1-829)

QY      1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db      578 TGGATGGGGGAGCTGCGACCGGGGCTGCCACCGCGCTGTGCGAGCATGGGACC 637

QY      21 CysArgAspGlyIysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db      638 TGCCCGGACGGCAAGTGGCGAGTGCAGCCCTGGCTGGAATGGCAACACTGCACCATCGCT 697

QY      41 HisTyrLeuAspArgValValIysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db      698 CACTATCTGGATAGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 757

QY      61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db      758 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCACTGCGTCTGCCACTGGGCTGGAGAGAGCTGGC 817

QY      81 Cys 81
```

```
Db      818 TGT 820

RESULT 13
US-10-383-201-45
; Sequence 45, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 45
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1392)
US-10-383-201-45

Alignment Scores:
Pred. No.:      6.97e-51      Length:      1392
Score:          509.00      Matches:      81
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    82.50%      Indels:      0
DB:             13          Gaps:          0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-45 (1-1392)

QY      1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db      1144 TGGATGGGGGAGCTGCGACCGGGGCTGCCACCGCGCTGTGCGAGCATGGGACC 1203

QY      21 CysArgAspGlyIysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db      1204 TGCCCGGACGGCAAGTGGCGAGTGCAGCCCTGGCTGGAATGGCAACACTGCACCATCGCT 1263

QY      41 HisTyrLeuAspArgValValIysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db      1264 CACTATCTGGATAGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 1323

QY      61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db      1324 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCACTGCGTCTGCCACTGGGCTGGAGAGAGCTGGC 1383

QY      81 Cys 81
Db      1384 TGT 1386
```


RESULT 14
US-10-383-201-53
; Sequence 53, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 53
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1386)
US-10-383-201-53

Alignment Scores:
Pred. No.: 6,97e-51 Length: 1392
Score: 509.00 Matches: 81
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-53 (1-1392)

| | | | |
|----|------|---|------|
| QY | 1 | TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr | 20 |
| Db | 1144 | TGGATGGGGCAGCCCTGCCAGCAGCGGCTGCCACCCGGCTGTGCCGAGCATGGGACC | 1203 |
| QY | 21 | CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla | 40 |
| Db | 1204 | TGCCCGCAGCGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACAACCTGCACCATCGCT | 1263 |
| QY | 41 | HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg | 60 |
| Db | 1264 | CACATCTGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCATGGCAACGGCAGA | 1323 |
| QY | 61 | CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly | 80 |
| Db | 1324 | TGTACCTTAGACCTGAATGTTGGCACTGCTGCCAGCTGGGCTGGAGAGGAGCTGGC | 1383 |
| QY | 81 | Cys 81 | |
| Db | 1384 | TGT | 1386 |

RESULT 15

US-10-383-201-41
; Sequence 41, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 41
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1464)
US-10-383-201-41

Alignment Scores:
Pred. No.: 7,47e-51 Length: 1476
Score: 509.00 Matches: 81
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 82.50%
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-41 (1-1476)

| | | | |
|----|------|---|------|
| QY | 1 | TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr | 20 |
| Db | 1222 | TGGATGGGGCAGCCCTGCCAGCAGCGGCTGCCACCCGGCTGTGCCGAGCATGGGACC | 1281 |
| QY | 21 | CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla | 40 |
| Db | 1282 | TGCCCGCAGCGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACAACCTGCACCATCGCT | 1341 |
| QY | 41 | HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg | 60 |
| Db | 1342 | CACATCTGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCATGGCAACGGCAGA | 1401 |
| QY | 61 | CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly | 80 |
| Db | 1402 | TGTACCTTAGACCTGAATGTTGGCACTGCTGCCAGCTGGGCTGGAGAGGAGCTGGC | 1461 |
| QY | 81 | Cys 81 | |
| Db | 1462 | TGT | 1464 |

Search completed: August 14, 2004, 19:17:13
Job time : 216.425 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 29,8807 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_750_850
Perfect score: 617
Sequence: 1 WNGAACDORACHPRCAHGT.....DTSMETACGDSKNDGDLV 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB-Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosu62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn 1.1 258 @runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------------|-------------------|
| 1 | 430.5 | 69.8 | 447 | US-09-833-381-1072 | Sequence 1072, Ap |
| 2 | 169.5 | 27.5 | 5561 | US-08-400-159-9 | Sequence 1, Appli |
| 3 | 169.5 | 27.5 | 5561 | US-08-611-729A-1 | Sequence 1, Appli |
| 4 | 169 | 27.4 | 13857 | US-09-620-312B-75 | Sequence 75, Appl |
| 5 | 160 | 25.9 | 6049 | US-08-793-273C-3 | Sequence 3, Appli |
| 6 | 160 | 25.9 | 6049 | PCT-US95-11684-3 | Sequence 3, Appli |
| 7 | 155 | 25.1 | 4208 | US-09-214-278-6 | Sequence 6, Appli |
| 8 | 155 | 25.1 | 4208 | US-09-068-740A-10 | Sequence 10, Appl |
| 9 | 155 | 25.1 | 4208 | US-09-855-722-6 | Sequence 6, Appli |
| 10 | 155 | 25.1 | 5458 | US-09-199-865-2 | Sequence 2, Appli |
| 11 | 155 | 25.1 | 5590 | US-08-882-046-1 | Sequence 1, Appli |
| 12 | 155 | 25.1 | 6464 | US-08-400-159-5 | Sequence 5, Appli |

| | | | | | | |
|----|-------|------|------|---|--------------------|-------------------|
| 13 | 155 | 25.1 | 6464 | 3 | US-08-611-729A-5 | Sequence 5, Appli |
| 14 | 154.5 | 25.0 | 3582 | 2 | US-08-400-159-9 | Sequence 9, Appli |
| 15 | 154.5 | 25.0 | 3582 | 3 | US-08-611-729A-9 | Sequence 9, Appli |
| 16 | 152.5 | 24.7 | 7286 | 4 | US-08-793-273C-1 | Sequence 1, Appli |
| 17 | 152.5 | 24.7 | 7286 | 5 | PCT-US95-11684-1 | Sequence 1, Appli |
| 18 | 146.5 | 23.7 | 1556 | 1 | US-08-597-545-3 | Sequence 3, Appli |
| 19 | 146.5 | 23.7 | 1556 | 1 | US-08-457-135-3 | Sequence 3, Appli |
| 20 | 146.5 | 23.7 | 2857 | 3 | US-09-981-392-4 | Sequence 4, Appli |
| 21 | 146 | 23.7 | 8091 | 4 | US-09-230-652-1 | Sequence 1, Appli |
| 22 | 146 | 23.7 | 8257 | 4 | US-09-484-970B-65 | Sequence 65, Appl |
| 23 | 144 | 23.3 | 716 | 4 | US-09-1423-753-11 | Sequence 11, Appl |
| 24 | 144 | 23.3 | 2508 | 3 | US-08-981-392-1 | Sequence 1, Appli |
| 25 | 144 | 23.3 | 2663 | 4 | US-08-088-740A-8 | Sequence 8, Appli |
| 26 | 144 | 23.3 | 2663 | 4 | US-09-423-753-8 | Sequence 8, Appli |
| 27 | 144 | 23.3 | 2883 | 3 | US-08-981-392-3 | Sequence 3, Appli |
| 28 | 143.5 | 23.3 | 4315 | 3 | US-08-882-046-3 | Sequence 3, Appli |
| 29 | 143.5 | 23.3 | 4464 | 2 | US-08-400-159-7 | Sequence 7, Appli |
| 30 | 143.5 | 23.3 | 4483 | 3 | US-08-611-729A-7 | Sequence 7, Appli |
| 31 | 142.5 | 23.1 | 1981 | 3 | US-08-981-392-26 | Sequence 26, Appl |
| 32 | 142.5 | 23.1 | 2005 | 4 | US-09-635-872A-18 | Sequence 18, Appl |
| 33 | 142.5 | 23.1 | 2005 | 4 | US-09-636-077A-18 | Sequence 18, Appl |
| 34 | 142.5 | 23.1 | 2005 | 4 | US-09-636-060C-18 | Sequence 18, Appl |
| 35 | 142.5 | 23.1 | 2005 | 4 | US-09-986-552-18 | Sequence 18, Appl |
| 36 | 142.5 | 23.1 | 4080 | 4 | US-09-016-434-1326 | Sequence 1326, Ap |
| 37 | 142.5 | 23.1 | 5645 | 4 | US-09-023-655-1319 | Sequence 1319, Ap |
| 38 | 142.5 | 23.1 | 5681 | 4 | US-09-919-172-58 | Sequence 58, Appl |
| 39 | 142.5 | 23.1 | 5933 | 4 | US-09-919-172-23 | Sequence 23, Appl |
| 40 | 141.5 | 22.9 | 2899 | 3 | US-08-981-392-24 | Sequence 24, Appl |
| 41 | 140.5 | 22.8 | 3955 | 3 | US-09-214-278-4 | Sequence 4, Appli |
| 42 | 140.5 | 22.8 | 3955 | 4 | US-09-855-722-4 | Sequence 4, Appli |
| 43 | 140 | 22.7 | 2206 | 4 | US-09-907-794A-3 | Sequence 3, Appli |
| 44 | 140 | 22.7 | 2206 | 4 | US-09-905-125A-3 | Sequence 3, Appli |
| 45 | 140 | 22.7 | 2206 | 4 | US-09-902-775A-3 | Sequence 3, Appli |

ALIGNMENTS

RESULT 1
US-09-833-381-1072
; Sequence 1072, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1072
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(447)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1072

Alignment Scores:
Pred. No.: 1.36e-36 Length: 447
Score: 430.50 Matches: 72
Percent Similarity: 88.89% Conservativeness: 0
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 69.77% Indels: 9
DB: 4 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x US-09-833-381-1072 (1-447)

QY 21 CysAtgAspGlyLysCysGluCysSerProGlyTyrAsnGlyGluHisCysThrIleAla 40

Db 23 TGGCGGACGCAAGTGGAGTGCAGCCCTGGTGGTGAATGGGCAACACTGCACCATC--- 79

Qy 41 HistyrlLeuAspArgValVallysgluGlyCysProglyLeuCysAsnGlyAsnGlyArg 60

Db 80 -----GAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 115

Qy 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80

Db 116 TGTACCTTAGACCTGAATGGTGGACATGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGCG 175

Qy 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100

Db 176 TGTGACATTCATGGAGACTGCTCGGTGCAGCAAGCAATGATGGAGATGGCCTG 235

Qy 101 Val 101

Db 236 GTG 238

RESULT 2

US-08-400-159-1

; Sequence 1, Application US/08400159

; Patent No. 5869282

; GENERAL INFORMATION:

; APPLICANT: Ish-Horowicz, David

; APPLICANT: Henrique, Domingos M.P.

; APPLICANT: Lewis, Julian H.

; APPLICANT: Myat, Anna M.

; APPLICANT: Fleming, Robert J.

; APPLICANT: Artavanis-Tsakonas, Spyridon

; APPLICANT: Mann, Robert S.

; APPLICANT: Gray, Grace E.

; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE

; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/400,159

; FILING DATE: 07-MAR-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Misrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 7326-029

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5561 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: unknown

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 442..4656

US-08-400-159-1

Alignment Scores: 6.48e-08 Length: 5561

Pred. No.: 169.50 Matches: 35

Score:

Percent Similarity: 45.16% Conservative: 7

Best Local Similarity: 37.63% Mismatches: 34

Query Match: 27.47% Indels: 17

DB: 2 Gaps: 6

US-10-029-020-14_COPY_750_850 (1-101) x US-08-400-159-1 (1-5561)

Qy 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCys---AlaGluHisGly 19

Db 1261 TGGCAGGGCGTCACTGCGAGAGGCCATATGCAAGGGGGCTGGACCCCGTCCACGGC 1320

Qy 20 ThrCys---ArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38

Db 1321 AAGTGGCATCTGCGGGGGAATGCGATGCAAGACCCGGCTGGCGTGTTCATTGTGCAAC 1380

Qy 39 IleAlaHistyrlLeuAspArgValVallysgluGlyCysPro---GlyLeuCysAsnGly 57

Db 1381 -----GAGTGCATGCTATCCCGGCTGCAAGCATGGTTCCTGCAACGGC 1425

Qy 58 AsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArg 77

Db 1426 -----AGCGCTGGAAATGGTGTGGCAGACCAACTGGGT 1461

Qy 78 GlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90

Db 1462 GGCATATTGTCGATCAAGATTTA---AATTTCGCGGC 1497

RESULT 3

US-08-611-729A-1

; Sequence 1, Application US/08611729A

; Patent No. 6004924

; GENERAL INFORMATION:

; APPLICANT: Ish-Horowicz, David

; APPLICANT: Henrique, Domingos M.P.

; APPLICANT: Lewis, Julian H.

; APPLICANT: Myat, Anna M.

; APPLICANT: Fleming, Robert J.

; APPLICANT: Artavanis-Tsakonas, Spyridon

; APPLICANT: Mann, Robert S.

; APPLICANT: Gray, Grace E.

; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE

; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/611,729A

; FILING DATE: 06-MAR-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Misrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 7326-037

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5561 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: unknown

```
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 442..4653
US-08-611-729A-1

Alignment Scores:
Pred. No.: 6.48e-08 Length: 5561
Score: 169.50 Matches: 35
Percent Similarity: 45.16% Conservative: 7
Best Local Similarity: 37.63% Mismatches: 34
Query Match: 27.47% Indels: 17
DB: 3 Gaps: 6

US-10-029-020-14_COPY_750_850 (1-101) x US-08-611-729A-1 (1-5561)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCys---AlaGluHisGly 19
Db 1261 TGGCAGGGCGTCAACTGCGAGGAGGCCATATGCAAGGCGGCTCGCAGCCCGTCCACGGC 1320

QY 20 ThrCys---ArgAspGlyLysCysGluCysSerProGlyTTPAsnGlyGluHisCysThr 38
Db 1321 AAGTGCATGTCGCGGGGAATGCAATGCAAGCCGCTGCGTGTGTCATTGTGCAAC 1380

QY 39 IleAlaHisTyrLeuAspArgValLysGluGlyCysPro---GlyLeuCysAsnGly 57
Db 1381 -----GAGTGCATGTCATCCCGGCTGCAAGCATGTTCTTCCACACGGC 1425

QY 58 AsnGlyArgCysThrLeuAspLeuAsnGlyTTPHisCysValCysGlnLeuGlyTTPArg 77
Db 1426 -----AGCGCCTGGAAATGCGTGTGCGACACCACTGGGGT 1461

QY 78 GlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1462 GGCATATTGGCATCAAGATTTA---AATTTCGCGGC 1497
```

```
RESULT 4
US-09-620-312D-75
; Sequence 75, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 75
; LENGTH: 13857
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..  
US-09-620-312D-75

Alignment Scores:
Pred. No.: 2.38e-07 Length: 13857
Score: 169.00 Matches: 35
Percent Similarity: 43.75% Conservative: 7
Best Local Similarity: 36.46% Mismatches: 34
Query Match: 27.39% Indels: 20
DB: 4 Gaps: 4

US-10-029-020-14_COPY_750_850 (1-101) x US-09-620-312D-75 (1-13857)

QY 3 GlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThrCysArg 22
Db 1555 GGTGAGGACTGTGGGAGCCGTCGCTGTCCCGGGAGCTGCCGTGGGACGGCCTTTGGCAG 1614

QY 23 AspGlyLysCysGluCysSerProGlyTTPAsnGlyGluHisCysThrIleAlaHisTyr 42
Db 1615 GATGGCGTGTGCGTGTGTGACGAGGCTACTCAGGGGAAGACTGCAGC----- 1662

QY 43 LeuAspArgValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArgCysThr 62
Db 1663 -----ACGCCAGCTGCCCGGGGCTGCCGAGGCCGCGCCAGTGC--- 1704

QY 63 LeuAspLeuAsnGlyTTPHisCysValCysGlnLeuGlyTTPArgGlyAlaGlyCys--- 81
Db 1705 -----CTAGATGGG---CGGTGTGTGTGGAGGACGGCTACTCTGGCGAGGATTGGCGT 1755

QY 82 -----AspThrSerMetGluThrAlaCysGlyAsp 91
Db 1756 GTGAGCAGTGTCCCGAATGACTGCAGCCAGCGGCTGTGCCAGGAC 1803
```

```
RESULT 5
US-08-793-273C-3
; Sequence 3, Application US/08793273C
; Patent No. 6482410
; GENERAL INFORMATION:
; APPLICANT: Crossin, Kathryn L.
; APPLICANT: Phillips, Greg
; APPLICANT: Prieto, Anne L.
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE ATTACHMENT AND
; FILE REFERENCE: BEC00225
; CURRENT APPLICATION NUMBER: US/08/793,273C
; CURRENT FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: PCT/US95/11684
; PRIOR FILING DATE: 1995-09-14
; PRIOR APPLICATION NUMBER: 08/308,359
; PRIOR FILING DATE: 1994-09-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 6049
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (309)..  
US-08-793-273C-3

Alignment Scores:
Pred. No.: 7.09e-07 Length: 6049
Score: 160.00 Matches: 33
Percent Similarity: 48.86% Conservative: 10
Best Local Similarity: 37.50% Mismatches: 31
Query Match: 25.93% Indels: 14
DB: 4 Gaps: 4

US-10-029-020-14_COPY_750_850 (1-101) x US-08-793-273C-3 (1-6049)
```

| | | | |
|----|------|--|------|
| QY | 1 | TripMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr | 20 |
| | | ::: ::: ::: | |
| Db | 1875 | TTTCATGGGGAGGACTGCGGGAGCTGTCTCTGCCAATGACTGCCACCACACGGGGCGC | 1934 |
| | | ::: ::: ::: | |
| QY | 21 | CysArgAspGlyLyCysGluCysSerProGlyTriPAsnGlyGluHisCysThrIleAla | 40 |
| | | ::: ::: ::: | |
| Db | 1935 | TGCGTCGATGGCGCGCTGGGTGTGCCAGAGGGCTTCACGTGGGAGAGACTGCCGG | 1988 |
| | | ::: ::: ::: | |
| QY | 41 | HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg | 60 |
| | | ::: ::: ::: | |
| Db | 1989 | -----GAACGGTCTCTGCCCAATGACTGCACAACTGGGGCGGC | 2027 |
| | | ::: ::: ::: | |
| QY | 61 | CysThrLeuAspLeuAsnGlyTriPHisCysValCysGlnLeuGlyTriPArgGlyAlaGly | 80 |
| | | ::: ::: ::: | |
| Db | 2028 | TGT-----GTGAGGGA---CGGTGTGTGTGTGAGGAGGTTACATGGGATCGAC | 2075 |
| | | ::: ::: ::: | |
| QY | 81 | Cys---AspThrSerMetGluThr | 87 |
| | | ::: ::: ::: | |
| Db | 2076 | TGTTCTCATGTGTCTCTCCCAACG | 2099 |
| | | ::: ::: ::: | |

RESULT 6
PCT-US95-11684-3
Sequence 3, Application PC/TUS9511684
GENERAL INFORMATION:
APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
TITLE OF INVENTION: CYTOSTATIN DERIVATIVES THAT STIMULATE
ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESS: Patent Counsel
STREET: 10666 North Torrey Pines Road, TPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11684
FILING DATE: 14-SEP-1995

| | | | |
|------------------------|----------|---------------|------|
| Alignment Scores: | 7.09e-07 | Length: | 6049 |
| Pred. No.: | 160.00 | Matches: | 33 |
| Score: | 48.86% | Conservative: | 10 |
| Percent Similarity: | 37.50% | Mismatches: | 31 |
| Best Local Similarity: | | | |

| | | | |
|---|--------|---|------|
| Query Match: | 25.93% | 14 | |
| DB: | 5 | 4 | |
| | | Indels: | |
| | | Gaps: | |
| US-10-029-020-14_COPY_750_850 (1-101) x PCT-US95-11684-3 (1-6049) | | | |
| Qy | 1 | TripMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr | 20 |
| | | ::: ::: ::: ::: ::: ::: ::: ::: | |
| Db | 1875 | TTCTATGGGGAGACTGCGGGAGCTCTCTCTGCCAATCACTGCCACCAGCAGCGGGCGC | 1934 |
| Qy | 21 | CysArgAspGlyIysCysGluCysSerProGlyTTPAsnGlyGluHisCysThrIleAla | 40 |
| | | ::: ::: ::: ::: ::: ::: ::: ::: | |
| Db | 1935 | TGGGTGATGGGGCGCTCGGTGCCACGAGGGCTTCACTGGGGAGAACTGCCGG----- | 1988 |
| Qy | 41 | HisTyrLeuAspArgValIysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg | 60 |
| | | ::: ::: ::: ::: ::: ::: ::: ::: | |
| Db | 1989 | -----GAACGGTCTGCCCAATGACTGCACAACTGGGGCGCGC | 2027 |
| Qy | 61 | CysThrIleuAspLeuAsnGlyTyrHisCysValCysGlnLeuGlyTyrArgGlyAlaGly | 80 |
| | | ::: ::: ::: ::: ::: ::: ::: ::: | |
| Db | 2028 | TGT-----GTCAGGGA---CGGTGTGTCTGTGAGGAAGTTACATGGGGATCGAC | 2075 |
| Qy | 81 | Cys---AspThrSerMetGluThr | 87 |
| | | ::: ::: ::: ::: ::: ::: ::: ::: | |
| Db | 2076 | TGTTCTGATGCTCTCTCTCCAAAGC | 2099 |

RESULT 7
US-09-214-278-6
; Sequence 6, Application US/09214278
; Patent No. 6291210
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/214,278
; CURRENT FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (409)..(4062)
; NAME/KEY: sig_peptide
; LOCATION: (409)..(501)
; NAME/KEY: mat_peptide
; LOCATION: (502)..(4062)
US-09-214-278-6

| Alignment Scores: | |
|------------------------|----------|
| Pred. No.: | 1,48e-06 |
| Score: | 155.00 |
| Percent Similarity: | 46.81% |
| Best Local Similarity: | 32.98% |
| Query Match: | 25.12% |
| DS: | 3 |
| Length: | 4208 |
| Matches: | 31 |
| Conservative: | 13 |
| Mismatch: | 13 |
| Indels: | 20 |
| Gaps: | 7 |

[illegible]

QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1240 -----GACCCCTGGCAGTGCCTCTGTGTGAGACCAACTGG 1272
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1273 GCGGCGCAGCTGTGTGACAAAGATCTC---AATTACTGTGGG 1311

RESULT 8

US-09-068-740A-10
; Sequence 10, Application US/09068740A
; Patent No. 6337387
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; APPLICANT: ITOH, AKIRA
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8447
; CURRENT APPLICATION NUMBER: US/09/068,740A
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: JP 7-299611
; PRIOR FILING DATE: 1995-11-17
; PRIOR APPLICATION NUMBER: JP 7-311811
; PRIOR FILING DATE: 1995-11-30
; PRIOR APPLICATION NUMBER: PCT/JP96/03356
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (409)...(4062)
US-09-068-740A-10

Alignment Scores:
Pred. No.: 1,48e-06 Length: 4208
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 7 Gaps: 4

US-10-029-020-14_COPY_750_850 (1-101) x US-09-068-740A-10 (1-4208)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 1078 TGGATGGGCGCCGGAATGTAACAGAGCTATTTCGCGACAAGGCTGCAGTCTTAAGCATGG 1137
QY 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
Db 1138 TCTTGCAAACTCCAGTGACTGCAGTGCCAGTACGGTGGCAAGGCTGTACTGT--- 1194
QY 39 IleAlaHisTyrLeuAspArgValValLysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 1195 -----GATAAGTGCATCCACACCCCGGATGCGTCCACGCACTCTGTAAT 1239
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1240 -----GAGCCCTGGCAGTGCCTCTGTGTGAGACCAACTGG 1272
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1273 GCGGCGCAGCTGTGTGACAAAGATCTC---AATTACTGTGGG 1311

RESULT 9

US-09-855-722-6
; Sequence 6, Application US/09855722
; Patent No. 6636741
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira

; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/855,722
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (409)...(4062)
; NAME/KEY: sig peptide
; LOCATION: (409)...(501)
; NAME/KEY: mat peptide
; LOCATION: (502)...(4062)
US-09-855-722-6

Alignment Scores:
Pred. No.: 1,48e-06 Length: 4208
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 7 Gaps: 4

US-10-029-020-14_COPY_750_850 (1-101) x US-09-855-722-6 (1-4208)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 1078 TGGATGGGCGCCGGAATGTAACAGAGCTATTTCGCGACAAGGCTGCAGTCTTAAGCATGG 1137
QY 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
Db 1138 TCTTGCAAACTCCAGTGACTGCAGTGCCAGTACGGTGGCAAGGCTGTACTGT--- 1194
QY 39 IleAlaHisTyrLeuAspArgValValLysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 1195 -----GATAAGTGCATCCACACCCCGGATGCGTCCACGCACTCTGTAAT 1239
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1240 -----GAGCCCTGGCAGTGCCTCTGTGTGAGACCAACTGG 1272
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1273 GCGGCGCAGCTGTGTGACAAAGATCTC---AATTACTGTGGG 1311

RESULT 10

US-09-199-865-2
; Sequence 2, Application US/09199865
; Patent No. 6433138
; GENERAL INFORMATION:
; APPLICANT: Zimin, Ann B.
; APPLICANT: Maciag, Thomas
; APPLICANT: Wong, Michael K.K.
; APPLICANT: Pepper, Michael S.
; APPLICANT: Montesano, Roberto
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC METHODS AND COMPOSITIONS
; FILE REFERENCE: 0036-101
; CURRENT APPLICATION NUMBER: US/09/199,865
; CURRENT FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: 60/018,841
; EARLIER FILING DATE: 1996-05-31
; EARLIER APPLICATION NUMBER: PCT/US97/09407
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2

```
; LENGTH: 5458
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-199-865-2

Alignment Scores:
Pred. No.: 2,07e-06 Length: 5458
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 4 Gaps: 7

US-10-029-020-14_COPY_750_850 (1-101) x US-09-199-865-2 (1-5458)

Qy 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 640 TGGATGGGCCCGCGAATGTAACAGAGCTATTTCGCCGACAAAGGCTGCAGTCCTTAAGCATGGG 699

Qy 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTyrAsnGlyGluHisCysThr 38
Db 700 TCTTGGCAAACTCCCGAGTGACTGCAGGTGCAGTACGGCTGCAGAGGCTGTACTGT--- 756

Qy 39 IleAlaHisTyrLeuAspArgValVallysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 757 -----GATAAGTGATCCACACCGGGATGGTCCACGGCATCTGTAT 801

Qy 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTyrHisCysValCysGlnLeuGlyTyr 76
Db 802 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 834

Qy 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 835 GCGCGCCAGCTCTGTGACAAAGATCTC---AATTACTGTGGG 873

RESULT 11
US-08-882-046-1
; Sequence 1, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1
; SEQUENCE CHARACTERISTICS
; LENGTH: 5590 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 414..4068
; US-08-882-046-1

Alignment Scores:
Pred. No.: 2,13e-06 Length: 5590
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 3 Gaps: 7

US-10-029-020-14_COPY_750_850 (1-101) x US-08-882-046-1 (1-5590)

Qy 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 1083 TGGATGGGCCCGCGAATGTAACAGAGCTATTTCGCCGACAAAGGCTGCAGTCCTTAAGCATGGG 1142

Qy 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTyrAsnGlyGluHisCysThr 38
Db 1143 TCTTGGCAAACTCCCGAGTGACTGCAGGTGCAGTATGGCTGCAGAGGCTGTACTGT--- 1199

Qy 39 IleAlaHisTyrLeuAspArgValVallysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 1200 -----GATAAGTGATCCACACCGGGATGGTCCACGGCATCTGTAT 1244

Qy 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTyrHisCysValCysGlnLeuGlyTyr 76
Db 1245 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 1277

Qy 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1278 GCGCGCCAGCTCTGTGACAAAGATCTC---AATTACTGTGGG 1316

RESULT 12
US-08-400-159-5
; Sequence 5, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,159
; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
```



```
;
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6464 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 371..4027
;
US-08-400-159-5
Alignment Scores:
Pred. No.: 2,57e-06 Length: 6464
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 2 Gaps: 7

US-10-029-020-14_COPY_750_850 (1-101) x US-08-400-159-5 (1-6464)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 1040 TGGATGGGCCCCGAATTAACAGAGCTATTTCGCCACAAGGCTGCAGTCCTAAAGCATGGG 1099
QY 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
Db 1100 TCTTGCAAATCCCAAGTGCAGTGCCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1156
QY 39 IleAlaHisTyrLeuAspArgValLysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 1157 -----GATAAGTGCATCCACACCCCGGATGCGTCCACGGCATCTGTAAT 1201
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1202 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 1234
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1235 GCGCGCCAGCTCTGTGACAAAGATCTC---AATTACTGTGGG 1273

RESULT 13
US-08-611-729A-5
; Sequence 5, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6464 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 371..4024
;
US-08-611-729A-5
Alignment Scores:
Pred. No.: 2,57e-06 Length: 6464
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 3 Gaps: 7

US-10-029-020-14_COPY_750_850 (1-101) x US-08-611-729A-5 (1-6464)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 1040 TGGATGGGCCCCGAATTAACAGAGCTATTTCGCCACAAGGCTGCAGTCCTAAAGCATGGG 1099
QY 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
Db 1100 TCTTGCAAATCCCAAGTGCAGTGCCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1156
QY 39 IleAlaHisTyrLeuAspArgValLysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 1157 -----GATAAGTGCATCCACACCCCGGATGCGTCCACGGCATCTGTAAT 1201
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1202 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 1234
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1235 GCGCGCCAGCTCTGTGACAAAGATCTC---AATTACTGTGGG 1273

RESULT 14
US-08-400-159-9
; Sequence 9, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; NUMBER OF SEQUENCES: 20
```

```

CORRESPONDENCE ADDRESS:
ADDRESSSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,159
FILING DATE: 07-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3582 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3582
US-08-400-159-9

Alignment Scores:
Pred. No.: 1.35e-06 Length: 3582
Score: 154.50 Matches: 33
Percent Similarity: 41.80% Conservative: 18
Best Local Similarity: 27.05% Mismatches: 40
Query Match: 25.04% Indels: 31
DB: 2 Gaps: 4

US-10-029-020-14_COPY_750_850 (1-101) x US-08-400-159-9 (1-3582)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgala-----CysHisProArgCysAlaGluHis 18
Db 1246 TGGTCTGGCCACAACTGTCATATAATATTAAATGATTGTCGTGGACAATGTCAGAAATGGA 1305
QY 19 GlyThrCysArgAsp-----GlyLysCysGluCysSerProGlyTrpAsnGly 34
Db 1306 GGATCTCTGCGGACTTGGTTAATGTTATCGGTGCATCTGTTCCACCTGGCTATGCAGGA 1365
QY 35 GluHisCysThrIleAlaHisTyrLeuAspArgValValLysGluGlyCysProGlyLeu 54
Db 1366 GATCACTGT-----GAGAAAGACATCAATGAATGTCAGTAACCTT 1407
QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
Db 1408 TCGATGAATGGGGTCACTGCCAGGATGAATCAATGATTCCAAATGTCATGTCCTGCT 1467
QY 75 GlyTrpArgGlyAlaGlyCys----- 81
Db 1468 GGTTCCTCAGGAACCTCTGTCAGCTGGATATAGACTACTGTGAGCCAAACCTTGGCCAG 1527
QY 82 -----AspThrSerMetGluThrAlaCysGlyAspSerLysAspAsn 95
Db 1528 AACGGTGCCAGTGGTTCATCTTGTCTATGGACTATTTCTGTAACTGCCCTGAAGATTAC 1587
QY 96 AspGly 97
Db 1588 GAAGGC 1593

RESULT 15
US-08-611-729A-9
; Sequence 9, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3579
US-08-611-729A-9

Alignment Scores:
Pred. No.: 1.35e-06 Length: 3582
Score: 154.50 Matches: 33
Percent Similarity: 41.80% Conservative: 18
Best Local Similarity: 27.05% Mismatches: 40
Query Match: 25.04% Indels: 31
DB: 3 Gaps: 4

US-10-029-020-14_COPY_750_850 (1-101) x US-08-611-729A-9 (1-3582)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgala-----CysHisProArgCysAlaGluHis 18
Db 1246 TGGTCTGGCCACAACTGTCATATAATATTAAATGATTGTCGTGGACAATGTCAGAAATGGA 1305
QY 19 GlyThrCysArgAsp-----GlyLysCysGluCysSerProGlyTrpAsnGly 34
Db 1306 GGATCTCTGCGGACTTGGTTAATGTTATCGGTGCATCTGTTCCACCTGGCTATGCAGGA 1365
QY 35 GluHisCysThrIleAlaHisTyrLeuAspArgValValLysGluGlyCysProGlyLeu 54
Db 1366 GATCACTGT-----GAGAAAGACATCAATGAATGTCAGTAACCTT 1407
QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
Db 1408 TCGATGAATGGGGTCACTGCCAGGATGAATCAATGATTCCAAATGTCATGTCCTGCT 1467
QY 75 GlyTrpArgGlyAlaGlyCys----- 81
Db 1468 GGTTCCTCAGGAACCTCTGTCAGCTGGATATAGACTACTGTGAGCCAAACCTTGGCCAG 1527
QY 82 -----AspThrSerMetGluThrAlaCysGlyAspSerLysAspAsn 95
Db 1528 AACGGTGCCAGTGGTTCATCTTGTCTATGGACTATTTCTGTAACTGCCCTGAAGATTAC 1587
QY 96 AspGly 97
Db 1588 GAAGGC 1593
```

```

Db 1366 GATCACTGT-----GAGAAAGACATCAATGAATGTGCAAGTAACCT 1407
Qy 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
Db 1408 TGCATGATGGGGTCACTGCGAGATGAAATCAATGGATTCCAATGCTGTGCTCTGCT 1467
Qy 75 GlyTrpArgGlyAlaGlyCys-----81
Db 1468 GGTTCCTCAGGAACCTCTGTGTCAGCTGGATATAGACTACTGTGAGCCAAACCTTGCCAG 1527
Qy 82 -----AspThrSerMetGluThrAlaCysGlyAspSerLysAspAsn 95
Db 1528 AACGGTGCCAGAGCTTCAATCTTGTCTATGACTATTCTGTAACGCGCTGAAGATTAC 1587
Qy 96 AspGly 97
Db 1588 GAAGGC 1593

```

Search completed: August 14, 2004, 19:40:00
Job time : 44.8807 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 1184.51 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_750_850
Perfect score: 617
Sequence: 1 WGAACDQACRACHPRCAEHGT.....DTSMETACGDSKNDGDGLV 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_13135 @runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_pro:*
26: em_gss_rod:*
27: em_gss_pbg:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|-----------|--------------------|
| 1 | 543.5 | 88.1 | 610 | 10 | BB633431 | BB633431 BB633431 |
| 2 | 543.5 | 88.1 | 1284 | 11 | AK039472 | AK039472 Mus muscu |
| 3 | 537.5 | 87.1 | 464 | 10 | BE671538 | BE671538 7653F01.x |
| 4 | 470.5 | 76.3 | 435 | 14 | CB793108 | CB793108 AMGNNUC:N |
| 5 | 454 | 73.6 | 462 | 10 | BE646762 | BE646762 UT-M-BH1- |
| 6 | 439 | 71.2 | 718 | 13 | BU341069 | BU341069 603406859 |
| 7 | 410.5 | 66.5 | 847 | 14 | CK017330 | CK017330 AGENCOURT |
| 8 | 323 | 52.4 | 314 | 9 | AL710634 | AL710634 DKFZp886G |
| 9 | 261.5 | 42.4 | 492 | 14 | CD163026 | CD163026 MLI-0081T |
| 10 | 254.5 | 41.2 | 501 | 14 | CD202234 | CD202234 MSI-0131U |
| 11 | 252.5 | 40.9 | 633 | 10 | BB665486 | BB665486 BB665486 |
| 12 | 252.5 | 40.9 | 683 | 10 | BB665259 | BB665259 BB665259 |
| 13 | 250 | 40.5 | 1376 | 11 | AK053790 | AK053790 Mus muscu |
| 14 | 250 | 40.5 | 3190 | 29 | AY413477 | AY413477 Mus muscu |
| 15 | 250 | 40.5 | 6246 | 29 | AY413475 | AY413475 Homo sapi |
| 16 | 247 | 40.0 | 2729 | 11 | AK031173 | AK031173 Mus muscu |
| 17 | 224 | 36.3 | 550 | 28 | AZ393761 | AZ393761 IM0156P20 |
| 18 | 223.5 | 36.2 | 511 | 10 | BE650761 | BE650761 UT-M-BH3- |
| 19 | 223.5 | 36.2 | 2627 | 11 | AK050784 | AK050784 Mus muscu |
| 20 | 216 | 35.0 | 1040 | 29 | CNS0087S | AL085915 Drosophi |
| 21 | 214.5 | 34.8 | 1101 | 29 | CNS00056 | AL050564 Drosophi |
| 22 | 210 | 34.0 | 252 | 10 | BE088902 | BE088902 CM2-BT069 |
| 23 | 210 | 34.0 | 252 | 10 | BE088973 | BE088973 CM2-BT069 |
| 24 | 206 | 33.4 | 1101 | 29 | CNS000GVU | AL072610 Drosophi |
| 25 | 202.5 | 32.8 | 376 | 14 | CD118097 | CD118097 ME1-0044T |
| 26 | 200 | 32.4 | 659 | 13 | BQ284539 | BQ284539 fz88405.x |
| 27 | 198 | 32.1 | 1023 | 29 | CNS006NW | AL085602 Drosophi |
| 28 | 171 | 27.7 | 724 | 13 | BU387957 | BU387957 603583120 |
| 29 | 171 | 27.7 | 2716 | 11 | AK031198 | AL167738 Tetraodon |
| 30 | 171 | 27.6 | 395 | 13 | BM188423 | AK031198 Mus muscu |
| 31 | 170.5 | 27.6 | 572 | 9 | AL588592 | BM188423 BM188423 |
| 32 | 169.5 | 27.5 | 589 | 12 | BI358014 | AL588592 AL588592 |
| 33 | 169.5 | 27.5 | 609 | 12 | BI356067 | BI358014 RE44666.5 |
| 34 | 169.5 | 27.5 | 733 | 13 | BU110388 | BI356067 RE42104.5 |
| 35 | 169.5 | 27.5 | 733 | 13 | BU110388 | BU110388 603002245 |
| 36 | 169.5 | 27.5 | 737 | 13 | BU325086 | BU325086 603492835 |
| 37 | 169.5 | 27.5 | 751 | 13 | BU274274 | BU325086 603492835 |
| 38 | 169.5 | 27.5 | 790 | 13 | BU273516 | BU274274 603530626 |
| 39 | 169.5 | 27.5 | 805 | 13 | BU284708 | BU273516 603533089 |
| 40 | 168.5 | 27.3 | 626 | 12 | BI393390 | BU284708 603601739 |
| 41 | 167.5 | 27.1 | 659 | 13 | BQ284142 | BI393390 pgp1n.pk0 |
| 42 | 167.5 | 27.1 | 805 | 13 | BU340855 | BQ284142 fz88405.y |
| 43 | 166.5 | 27.0 | 611 | 13 | BU462653 | BU340855 603518278 |
| 44 | 165.5 | 26.8 | 675 | 13 | BU403903 | BU462653 603366851 |
| 45 | 163 | 26.4 | 755 | 13 | BU408094 | BU403903 604138914 |

ALIGNMENTS

RESULT 1
BB633431
LOCUS
DEFINITION BB633431 RIKEN full-length enriched, adult male spinal cord Mus
musculus cDNA clone A330048C04 5', mRNA sequence.
610 bp mRNA linear EST 26-OCT-2001
ACCESSION BB633431.1 GI:16469876
VERSION BB633431
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 610)

AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y. RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

TITLE

JOURNAL

COMMENT

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suihoro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome.* 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

e mouse tissues.

FEATURES

source

```
Location/Qualifiers
1..610
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A330048C04"
/sex="male"
/tissue_type="spinal cord"
/lab_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male spinal cord"
```

/note="Site 1: Sall; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATTCGAGTAAATAATCCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Alignment Scores:

Pred. No.: 2.5e-37 Length: 610
Score: 543.50 Matches: 89
Percent Similarity: 88.12% Conservative: 0
Best Local Similarity: 88.12% Mismatches: 3
Query Match: 88.09% Indels: 9
DB: 10 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x BB633431 (1-610)

QY 1 TtpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20

Db 167 TGGATGGGGCGCGCATGCGACCAACGCGGCTGCCACCCACCGCTGTGCAGAACACGGGACC 226

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTtpAsnGlyGluHisCysThrIleAla 40

Db 227 TGCCGGGACGGCAAGTGCAGATGCAGCCCGCTGGAAATGGAGAGCACTGCACCATC--- 283

QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyValAsnGlyArg 60

Db 284 -----GAGGGCTGTCTCTGCTGTGCAATGGAAATGGCAGA 319

QY 61 CysThrLeuAspLeuAsnGlyTtpHisCysValCysGlnLeuGlyTtpArgGlyAlaGly 80

Db 320 TGTACCTGGACCTGAATGGGTGGCACTGTGTCTGCCAGCTGGGCTGGCGAGGAGACTGGC 379

QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnGlyValAspGlyLeu 100

Db 380 TGCACATCATCATGAAACGGCTGTGGAGCGCAAGGCAACACGCGAGATGCGCTTG 439

QY 101 Val 101

Db 440 GTG 442

RESULT 2

AK039472

LOCUS

DEFINITION

Mus musculus adult male spinal cord cDNA, RIKEN full-length enriched library, clone:A330048C04 product:neuregulin 1, full insert sequence.

AK039472 1284 bp mRNA linear HTC 19-SEP-2003

AK039472.1 GI:26087188

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

```

MEDLINE
PUBMED
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1284)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakanura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,K., Saito,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muranatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source
1. .1284
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A330048C04"
/db_xref="MGI:2403942"
/db_xref="taxon:10090"
/clone="A330048C04"
/sex="male"
/tissue_type="spinal cord"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1. .1284
/note="neuregulin 1 (MGD|MGI:96083, MGD|MGI:1345182,
GB|NW_011858, evidence: BLASTN, 99%, match=186)"
ORIGIN
Alignment Scores:
Pred. No.: 7 03e-37 Length: 1284
Score: 543.50 Matches: 89
Percent Similarity: 88.12% Conservative: 0
Best Local Similarity: 88.12% Mismatches: 3
Query Match: 88.09% Indels: 9
DB: 11 Gaps: 1
US-10-029-020-14_COPY_750_850 (1-101) x AK039472 (1-1284)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db TGGATGGGGCCGATCGGACCAACGGCTGCCACCCCGCTGTGCAGACACGGGACC 226

```

```

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db TGGCGGGACGGCAAGTGCGAATGCGCCCGGTGGAATGAGAGCACTGCACCATC--- 283
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db -----GAGGGCTGCTCGCTGTGCAATGGAATGGGAGA 319
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyValaGly 80
Db TGTACCTCGACCTGAATGGTGTGCTGTGCTGCAGCTGGCTGGCGAGGACTGGC 379
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db TGGCGACACATCCATGGAACGGCTGTGGAGACGGCAAGCAACGACGAGGATGGCTTG 439
QY 101 Val 101
Db GTG 442
RESULT 3
BE671538
LOCUS
DEFINITION
7e53f01.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3286201 3'
similar to TR:070465 O70465 LOC4. ? mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
REFERENCE
1 (bases 1 to 464)
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/INL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco.
FEATURES
source
Location/Qualifiers
1. .464
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3286201"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu24"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; plasmid DNA from the normalized
library NCI CGAP Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneIDs
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."
ORIGIN
Alignment Scores:
Pred. No.: 5.67e-37 Length: 464
Score: 537.50 Matches: 88
Percent Similarity: 89.80% Conservative: 0

```

```

Best Local Similarity: 89.80%      Mismatches: 1
Query Match: 87.12%      Indels: 9
DB: 10      Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x BB671538 (1-464)

Qy 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 197 TGGATGGGGGAGCCCTGGACAGAGGGCCCTGCCACCCCGCTGGCCGAGATGGGACC 256

Qy 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 257 TGCCGCGACGCAAGTCGAGTCGAGCCCTGCTGGAATGGCGAACACTGCGACCATC--- 313

Qy 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 314 -----GAGGGTTGCCCTGGCTGGTGTGCAATGGCAACGCGAGA 349

Qy 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 350 TGTACCTTAGACCTGAATGGTTGGACATCGCTCTCCAGCTGGGCTGGAGGAGCTGGC 409

Qy 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAsp 98
Db 410 TGTGACACTTCCATGAGACTGCCTGGCTGACCAAGAAAGACAATGATGGAGAT 463

RESULT 4
CB793108
LOCUS CB793108 435 bp mRNA linear EST 16-MAY-2003
DEFINITION AMGNNUC:NRH5-00339-E12-A W Rat hypothalamus (10471) Rattus
norvegicus cDNA clone nrh5-00339-e12 5', mRNA sequence.
ACCESSION CB793108
VERSION CB793108.1 GI:29881501
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 435)
AUTHORS Angen EST Program.
TITLE Angen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00339 row: e column: 12.
FEATURES
Location/Qualifiers
1..435
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrh5-00339-e12"
/clone_lib="W Rat hypothalamus (10471)"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; W Rat
hypothalamus adult female Wistar rat avg. Insert size 2.3
kb fraction 6 and 7"

ORIGIN
Alignment Scores:
Pred. No.: 3.28e-31 Length: 435
Score: 470.50 Matches: 73
Percent Similarity: 83.00% Conservative: 10
Best Local Similarity: 73.00% Mismatches: 16
Query Match: 76.26% Indels: 1
DB: 14 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x CB793108 (1-435)

Qy 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 60 TGGGTAGGACCAACATGTGAGGACGTTCCCTGTCACCTCACTGCTGAGCATGCTGAC 119

```

```

Qy 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 120 TGCAGACGCGAAATGTAGTGTAGCCCTGGATGGAGGGCGACCATCTGCACAAATGCT 179

Qy 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 180 CACTACTTAGAT---GCTGTTGAGATGCTGTCCAGGGCTCTGCTTTGGAAATGGACGA 236

Qy 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 237 TGTACCTTGGATCAAAATGGTTGGCACTGTGTGTGTCAGGTGGTGGTGGAGTGGACAGGT 296

Qy 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 297 TGCACATTCATGCGAAATGCTTTGTGGCACAACACTTGGACATGATGAGATGCTTTG 356

```

```

RESULT 5
BB646762 462 bp mRNA linear EST 06-SEP-2000
LOCUS UI-M-BH1-akr-d-12-0-UI.r1 NIH_BMAP_M_S2 Mus musculus cDNA clone
DEFINITION UI-M-BH1-akr-d-12-0-UI 5', mRNA sequence.
ACCESSION BE646762
VERSION BE646762.1 GI:9972517
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 462)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 889548
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.

```

FEATURES

```

Location/Qualifiers
1..462
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH1-akr-d-12-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH_BMAP_M_S2"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH_BMAP_M_S2 library is a subtracted library derived from
NIH_BMAP_M_S1, which in turn is a subtracted library
derived from a mixture of normalized libraries from ten
regions of the mouse brain (cerebellum, brain stems,
olfactory bulbs, hypothalamus, cortex, amygdala, basal
ganglia, pineal gland, striatum, hippocampus). The driver
used for subtraction consisted of a pool of 5,000 clones
from the NIH_BMAP_M_S1 library and a pool of 2,000 clones
obtained from non-normalized and normalized mouse brain
spinal cord libraries."

```

ORIGIN

Alignment Scores: 9.58e-30 Length: 462
 Pred. No.: 454.00 Matches: 71
 Score: 77.23% Conservative: 7
 Percent Similarity: 70.30% Mismatches: 23
 Best Local Similarity: 73.58% Indels: 0
 Query Match: 10 Gaps: 0
 DB:

US-10-029-020-14_COPY_750_850 (1-101) x BE646762 (1-462)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
 Db 21 TGGACAGGCGCAGCTTGTGACACGCGGTGTGCCACCCCGCTGCATTGAGCAGCGGACC 80
 QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
 Db 81 TGTAAAGATGGCAAAATGTGAATCCGAGAGGGCTGGAATGGTGAACACTGCACCATTTGGT 140
 QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
 Db 141 AGCAACAGCGCAGGCGCCGAAACAGATGGCTGCCCTGATTGTGCAACGGTAACGGGAGA 200
 QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
 Db 201 TGCACACTGGGTGAGAACAGCTGGCAGTGTGTCTGCCAGACCGCTGGAGAGGGCCTGGA 260
 QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
 Db 261 TGCACAGTTGCCATGGAAACCTCCTCGCTGTATACAAAGGATAATGAGGAGATGCGCTG 320
 QY 101 Val 101
 Db 321 GTG 323

RESULT 6
 BU341069 718 bp mRNA linear EST 28-NOV-2002
 LOCUS 603406859F1 CSEQCHN67 Gallus gallus cDNA clone CHEST315el6 5', mRNA
 DEFINITION sequence.
 ACCESSION BU341069
 VERSION BU341069.1 GI:25849070
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 718)
 Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 12445392
 Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1. .718
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, HiseX"
 /db_xref="taxon:9031"
 /clone="CHEST315el6"
 /dev_stage="16 day embryo"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN67"

FEATURES
 source
 Location/Qualifiers
 1. .718
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, HiseX"
 /db_xref="taxon:9031"
 /clone="CHEST315el6"
 /dev_stage="16 day embryo"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN67"

/note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
 EcoRI; Site_2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN

Alignment Scores: 3.51e-28 Length: 718
 Pred. No.: 439.00 Matches: 70
 Score: 75.49% Conservative: 7
 Percent Similarity: 68.63% Mismatches: 24
 Best Local Similarity: 71.15% Indels: 1
 Query Match: 13 Gaps: 0
 DB:

US-10-029-020-14_COPY_750_850 (1-101) x BU341069 (1-718)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
 Db 31 TGGACAGGAGTGGCTGTGACACGCTGTGTGTATCCCGGTGTACAGACGCGAACT 90
 QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
 Db 91 TGTAAAGATGGGAAATGTGAATGACAGAGGGCTGGAATGGGGAGCACTGCACCATTTGGT 150
 QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
 Db 151 AGCAACAGCAGACGCCACCGAACAGATGGCTGCCCTGACTTGTGCATGGCAACGGGAGG 210
 QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
 Db 211 TGCACGCTGGCGCAGAACAGCTGGCAGTGTGTCTGCCAGACCGCTGGAGAGGCGCTGGA 270
 QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLe 100
 Db 271 TGCACAGTTGCCATGGAAACCTCCTGTGCCGATACACAGGATAAACGAGGAGATGCGCTT 330
 QY 100 uVal 101
 Db 331 GGT 334

RESULT 7

CK017330 847 bp mRNA linear EST 26-NOV-2003
 AGENCOURT_16543817 NIH_ZGC_10 Danio rerio cDNA clone IMAGE:704835
 5', mRNA sequence.
 LOCUS CK017330
 DEFINITION CK017330.1 GI:38543254
 ACCESSION EST.
 VERSION CK017330.1
 KEYWORDS Danio rerio (zebrafish)
 SOURCE Danio rerio
 ORGANISM Danio rerio

REFERENCE

1 (bases 1 to 847)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Len Zon, Harvard

REFERENCE

1 (bases 1 to 847)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Len Zon, Harvard

CDNA Library Preparation: Open Biosystems
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LAM14806 row: 1 column: 17
 High quality sequence stop: 585.

FEATURES

Location/Qualifiers
 1..847

/organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:7044835"
 /tissue_type="whole body"
 /lab_host="DH10B"
 /clone_lib="NIH ZGC 10"
 /note="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
 Bulk tissue was collected from a whole adult individual
 from the Tuebingen strain. 1st strand cDNA was primed with
 a Not I - oligo(dT) primer, double-stranded cDNA was
 cloned into the Not I and EcoRV sites of pExpress-1.
 Library was size-selected for >1 kb fragments. A
 normalized version of this library is also available
 (NIH ZGC 7). Library was constructed by Open Biosystems
 (Huntsville, AL)."

ORIGIN

Alignment Scores:
 Pred. No.: 1.3e-25 Length: 847
 Score: 410.50 Matches: 64
 Percent Similarity: 71.29% Conservative: 8
 Best Local Similarity: 63.37% Mismatches: 20
 Query Match: 66.53% Indels: 9
 DB: 14 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x CK017330 (1-847)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
 Db 209 TGGAGGGAACCACTGTGTGACAAACAGCCCTGCCATCAATCTGGAAGAGCATGGAGAA 268
 QY 21 CysArgAspGlyLeuGluCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
 Db 269 TGCAGGAGCGGCGAGTGTCTGTGTGAGCTGTGTGGAGGGAGAGCACTGCATATT--- 325
 QY 41 HisTyrLeuAspArgValValysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
 Db 326 -----GATGCTGCCCGAGGGTTGTGTAACGGTAATGGAGCGC 361
 QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
 Db 362 TGCACGCTGGAGCAGAGCGGTGGCATTTGTGTGCCAGTCTGGATGGAGCGGGCGGCGC 421
 QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyLeu 100
 Db 422 TGTAACGTCGTCATGGAACCGAGTGCATGACAGTAAAGATAATGACAGCGATGGTCTC 481
 QY 101 Val 101
 Db 482 ATG 484

RESULT 8
 AL710634
 LOCUS
 DEFINITION DKFZp686G1170_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
 DKFZp686G1170_5', mRNA sequence.
 AL710634
 ACCESSION
 VERSION AL710634.1 GI:19693989
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

1 (bases 1 to 314)
 Ansoorge, W., Wirkner, U., Mewes, W., Weil, B. and Wiemann, S.
 EST (Ansoorge, W., Wirkner, U., Mewes, W., Weil, B. and Wiemann, S.)
 Unpublished (1999)
 Contact: MIPS
 MIPS
 Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 Heidelberg/Germany within the cDNA sequencing consortium of the
 German Genome Project.
 No sl sequence available.
 This clone (DKFZp686G1170) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 Location/Qualifiers
 1..314
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686G1170"
 /dev stage="adult"
 /lab_host="DH10B"
 /clone_lib="686 (synonym: hlcc3)"
 /note="Vector: pTriplex2, Site 1: SfiI; Site 2: SfiI; B;
 cDNA-collection"

FEATURES

source
 1..314
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686G1170"
 /dev stage="adult"
 /lab_host="DH10B"
 /clone_lib="686 (synonym: hlcc3)"
 /note="Vector: pTriplex2, Site 1: SfiI; Site 2: SfiI; B;
 cDNA-collection"

ORIGIN

Alignment Scores:
 Pred. No.: 1.24e-18 Length: 314
 Score: 223.00 Matches: 51
 Percent Similarity: 72.60% Conservative: 2
 Best Local Similarity: 69.86% Mismatches: 20
 Query Match: 52.35% Indels: 0
 DB: 9 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x AL710634 (1-314)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
 Db 90 TGGACAGCGCAGCGGTGTGACAGCGCGTGTGCCACCCCGCTGATTGACAGCGGACC 149
 QY 21 CysArgAspGlyLeuGluCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
 Db 150 TGTAAAGATGGCAATGTGAATGCCGAGAGGCTGGATGGTGAACACTGCACCATTTGGT 209
 QY 41 HisTyrLeuAspArgValValysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
 Db 210 AGGCAACCGCGAGCGCACCGAAACAGATGCTGCCCTGACTTGTGCAACGGTAACGGGAGA 269
 QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGln 73
 Db 270 TGCACACTGGGTTCAGAACACAGCTGGCAGTGTGTCTGGCAA 308

RESULT 9

CD163026
 LOCUS
 DEFINITION ML1-0081T-M291-B11-U.G ML1-0081 Schistosoma mansoni cDNA clone
 ML1-0081T-M291-B11.G, mRNA sequence.

ACCESSION
 VERSION CD163026

KEYWORDS
 EST.

SOURCE
 Schistosoma mansoni

ORGANISM
 Schistosoma mansoni

REFERENCE
 1 (bases 1 to 492)

AUTHORS
 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
 Strigeida; Schistosomatoidea; Schistosomatidae; Schistosoma.

Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M.,
 Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,
 Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,


```

Qy 79 AlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLyAspAsnAspGlyAsp 98
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 327 AGTCGGTGTCAACGACAAATCGAACAATATGTGATGATGGAATAGATGATCATGAT 386
   |||||
Qy 99 GlyLeuVal 101
   |||||
Db 387 GATCTTGTT 395
   |||||

RESULT 11
BB665486 633 bp mRNA linear EST 26-OCT-2001
LOCUS BB665486 RIKEN full-length enriched, 0 day neonate eyeball Mus
DEFINITION musculus cDNA clone E130308J18 5', mRNA sequence.
ACCESSION BB665486
VERSION BB665486.1 GI:16499119
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 633)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamana,K.I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. .633
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="E130308J18"
/tissue_type="eyeball"
/dev_stage="0 day neonate"

FEATURES
source
/lab host="DH10B"
/clone lib="RIKEN full-length enriched, 0 day neonate
eyeball"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGCGCGCCCACTGCGAGTTTCTTTTCTTTTNN 3']. cDNA was
prepared by using thermose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATCTCGAGTTCTTAATAATATCCCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."
ORIGIN
Alignment Scores: 4.16e-12 Length: 633
Pred. No.: 252.50 Matches: 48
Score: 69.23% Conservative: 6
Percent Similarity: 61.54% Mismatches: 15
Best Local Similarity: 40.92% Indels: 9
Query Match: 10 Gaps: 3
DB:
US-10-029-020-14_COPY_750_850 (1-101) x BB665486 (1-633)
Qy 1 TrpMetGlyValAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
   |||||
Db 193 TGGATGGGGCGCGATGCGACCAACGGGCGCTGCCACCCAGCGCTGTGCAGAACACGGGACC 242
   |||||
Qy 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
   |||||
Db 243 TGCCGGGACGGCAGTGGCAATGCGAGCCCGCGCTGGAATGGAGAGCACTGCACCATCGGT 302
   |||||
Qy 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
   |||||
Db 303 ATG-----GCAGGGCGACGCTCTGGGCGCATTTGGGTGCTCTCAGAGACATTTCCAG 353
   |||||
Qy 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGly 78
   |||||
Db 354 GGCACCGCTGTTCTC-----CTCCATTGTCATG-----TCGATTGGC 389
   |||||

RESULT 12
BB665259 683 bp mRNA linear EST 31-AUG-2001
LOCUS BB665259 RIKEN full-length enriched, 0 day neonate eyeball Mus
DEFINITION musculus cDNA clone E130115C05 5', mRNA sequence.
ACCESSION BB665259
VERSION BB665259.1 GI:15403119
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 683)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamana,K.I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. .633
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="E130308J18"
/tissue_type="eyeball"
/dev_stage="0 day neonate"

```

Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A.,
Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K.,
Arakawa, T., Ishii, Y. and Hayashizaki, Y.

Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
Func. Genomics 2 pre, L72-L86 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

e mouse tissues.

Location/Qualifiers

1. .683
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="E130115C05"
/tissue_type="eyeball"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 0 day neonate
eyeball"

/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGCGCGCACTCGAGTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATTCGAGTTAATTAATTAATCCCGCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Alignment Scores:
Pred. No.: 4,62e-12 Length: 683
Score: 252.50 Matches: 48
Percent Similarity: 69.23% Conservative: 6
Best Local Similarity: 61.54% Mismatches: 15
Query Match: 40.92% Indels: 9
DB: 10 Gaps: 3

US-10-029-020-14_COPY_750_850 (1-101) x BB665259 (1-683)

QY 1 TtpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||||
Db 176 TGGATGGGGCCGCATGCGACCAACGGGCGTGCACCCGCGTGCAGAACACGGGACC 235
|||||
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTtpAsnGlyGluHisCysThrIleAla 40
|||||

Db 236 TGGCGGAGCGCAAGTGGGAATGAGCCCCCGGTGGGAATGAGAGCACTGCACCATCGGT 295
QY 41 HistYrLeuAspArgValLysGluGlyCysProGlyLeuCysAsnGlyArg 60
|||||
Db 296 ATG-----GCGGGCGCAGCGTCTGGGGCACTTGGTGGTCTCTCAGACATTCACG 346
QY 61 CysThrLeuAspLeuAsnGlyTtpHisCysValCysGlnLeuGlyTtpArgGly 78
|||||
Db 347 GCGAGCGCTGGTCTCTC-----CTCCATTGCATG-----TGGATTGGC 382

RESULT 13

AK053790

LOCUS

AK053790

DEFINITION

Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched

library, clone:E130308J18 product:neuregulin 1, full insert

sequence.

ACCSSION

AK053790

VERSION

AK053790.1 GI:26095991

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

AUTHORS

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

MEDLINE

99279253

PUBMED

10349636

REFERENCE

2

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

MEDLINE

20493174

PUBMED

11042159

REFERENCE

3

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL

MEDLINE

20530913

PUBMED

11076861

REFERENCE

4

AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the

PANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

JOURNAL

REFERENCE

5

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

JOURNAL

REFERENCE

6

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,

Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Konda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,

Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Takahashi, F., Takaku-Akahira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,

2 (bases 1 to 3190)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 Location/Qualifiers
 1..3190
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>3190
 /locus_tag="HCM4903"

ORIGIN
 Alignment Scores: 6.45e-11 Length: 3190
 Pred. No.: 250.00 Matches: 39
 Score: 100.00% Conservat: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 40.52% Gaps: 0
 DB: 29

US-10-029-020-14_COPY_750_850 (1-101) x AY413477 (1-3190)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
 226 TGATGGGGGGCGCATGTCAGCAACAGCGCCCTGCCACCACCGCTGTGCAGAACACGGACC 285
 QY 21 CysArgAspGlyLYbCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle 39
 286 TGC CGGAGCGCAAGTGGATGTCAGCCCGCTGGAATGGAGAGCACTGCACCATC 342

RESULT 15
 AY413475
 LOCUS
 DEFINITION
 Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 AY413475
 AY413475.1 GI:39769437
 GSS.
 SOURCE
 Homo sapiens (human)
 Organism
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 6246)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 Location/Qualifiers
 1..6246
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

gene <1..>6246
/locus_tag="HCM4903"

ORIGIN

Alignment Scores:
Pred. No.: 1.64e-10 Length: 6246
Score: 250.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.52% Indels: 0
DB: 29 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x AY413475 (1-6246)

| | | | |
|----|-----|--|-----|
| QY | 1 | TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr | 20 |
| Db | 226 | TGGATGGGGGACGCTCGACCCAGCGGGCTTGCACCCCGGCTGTGCCGAGCATGGGACC | 285 |
| QY | 21 | CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle | 39 |
| Db | 286 | TGCCCGGACGGCAAGTCCGAGTGCAGCCCTGGCTGGATGGCGACACTGCACCATC | 342 |

Search completed: August 14, 2004, 18:02:56
Job time : 1193.51 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 171.323 Seconds

(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_750_850

Perfect score: 617

Sequence: 1 WNGAACDQACPRCAEHCT.....DTSMETACGSKNDGDGLV 101

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q/cgn2_1/USPTO.spool/US10029020/runat_06082004_112215.29265/app.query.fasta_1.3519
-DB=N Geneseq 29Jan04 -QFMT=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1868@runat_06082004_112215.29265 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|-------------|
| 1 | 617 | 100.0 | 8354 | ABSS2100 | Human TEN |
| 2 | 617 | 100.0 | 8645 | ABSS78652 | Human cDN |
| 3 | 579 | 93.8 | 8438 | ABN85378 | Human NOV |
| 4 | 515 | 83.5 | 8473 | ABQ82345 | Human NOV |
| 5 | 515 | 83.5 | 8645 | ABQ82344 | Human NOV |
| 6 | 515 | 83.5 | 8675 | ABQ82343 | Human NOV |
| 7 | 478.5 | 77.6 | 8487 | ABQ82346 | Human NOV |
| 8 | 464.5 | 75.3 | 12879 | ABK92230 | Prostate |

| | | | | | |
|----|-------|------|--------|---|----------|
| 9 | 464.5 | 75.3 | 13202 | 4 | AAK51828 |
| 10 | 454 | 73.6 | 9826 | 5 | AAK51828 |
| 11 | 454 | 73.6 | 9826 | 5 | ADB32023 |
| 12 | 448.5 | 72.7 | 9729 | 5 | AAK51828 |
| 13 | 448.5 | 72.7 | 9729 | 5 | ADB32028 |
| 14 | 448.5 | 72.7 | 9729 | 5 | ADB32028 |
| 15 | 330.5 | 53.6 | 9058 | 7 | AAU60066 |
| 16 | 330.5 | 53.6 | 9058 | 7 | ACCT2051 |
| 17 | 296 | 48.0 | 10242 | 7 | ACCT2052 |
| 18 | 278 | 45.1 | 17131 | 4 | ABL29075 |
| 19 | 277.5 | 45.0 | 2569 | 4 | ABL17563 |
| 20 | 169.5 | 27.5 | 5552 | 4 | ABL10363 |
| 21 | 169.5 | 27.5 | 5561 | 3 | AAZ49039 |
| 22 | 169 | 27.4 | 13268 | 6 | ABV94056 |
| 23 | 169 | 27.4 | 13857 | 4 | AAI58199 |
| 24 | 169 | 27.4 | 13857 | 8 | ADB48165 |
| 25 | 169 | 27.4 | 100267 | 6 | ABT11032 |
| 26 | 160 | 25.9 | 6049 | 2 | AAI14548 |
| 27 | 159.5 | 25.9 | 5561 | 2 | AAQ43910 |
| 28 | 156 | 25.3 | 5353 | 8 | AAU62049 |
| 29 | 156 | 25.3 | 6696 | 8 | AAU62049 |
| 30 | 156 | 25.3 | 7560 | 4 | AAU62049 |
| 31 | 156 | 25.3 | 7560 | 6 | ABL62483 |
| 32 | 156 | 25.3 | 7560 | 7 | ABZ77285 |
| 33 | 156 | 25.3 | 7560 | 9 | ABZ77285 |
| 34 | 156 | 25.3 | 7560 | 9 | ACF79935 |
| 35 | 156 | 25.3 | 7746 | 6 | ABSS1834 |
| 36 | 156 | 25.3 | 7823 | 7 | ACA04023 |
| 37 | 156 | 25.3 | 7823 | 7 | ABX63790 |
| 38 | 155 | 25.1 | 4208 | 2 | AAI70175 |
| 39 | 155 | 25.1 | 4208 | 2 | AAV15201 |
| 40 | 155 | 25.1 | 4333 | 6 | ACD06197 |
| 41 | 155 | 25.1 | 4855 | 6 | ABK35564 |
| 42 | 155 | 25.1 | 5458 | 2 | AAV03674 |
| 43 | 155 | 25.1 | 5590 | 2 | AAV63753 |
| 44 | 155 | 25.1 | 5590 | 8 | ACD06196 |
| 45 | 155 | 25.1 | 5896 | 7 | ABZ34802 |

ALIGNMENTS

RESULT 1

ABSS2100

ID ABSS2100 standard; DNA; 8354 BP.

XX AC ABSS2100;

XX DT

XX 05-NOV-2002 (first entry)

XX DE

XX Human TEN-M4-like gene.

XX KW

Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis; cell signal processing; metabolic pathway modulation; metabolic disorder; obesity; diabetes; infectious disease; neurodegenerative disorder; acne; Alzheimer's disease; Parkinson's disease; immune disorder; cancer; haematopoietic disorder; cirrhosis; pancreatitis; learning defect; memory defect; infertility; congenital heart defect; hair growth; pigmentation disorder; endocrine disorder; respiratory disease; health; gastro-intestinal disease; reproductive; neurological disease; bone marrow transplantation; endocrine disease; allergy; inflammation; neurological disorder; urinary system disease; age-related disorder; neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4; adipocyte complement-related C1q tumour necrosis factor; out at first; beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter; type Ia membrane sushi-containing domain; butyrophilin; single nucleotide polymorphism.

OS Homo sapiens.

XX KW

XX Key

XX variation

XX Location/Qualifiers

XX replace(117,G)

XX /*tag= a

FT


```
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2249 TGGACGGGCCAGCCTGTAATCAGAGAGCCTGCCACCCCGCTGTGCCGAGCAGGGACC 2308

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle--- 39
Db 2309 TGGCCGACGCGCAAGTGGAGTGCAGCCCTGGCTGGAAATGGCGAACACTGCACCATCTCC 2368

QY 40 ---AlaHisTyrLeuAspArgValLys-----GluGlyCysProGlyLeuCysAsn 56
Db 2369 CTAGCTCACTATCTGGATAGGTAGTTAAACTTTTCAGAGGGTTCCTCGGTGTGTGCAAT 2428

QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 2429 GGCAACGGCAGATGACTTACCTAGACTGAATGGTTGGCACTCGCTGCCACTGGGCTGG 2488

QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAsp 96
Db 2489 AGAGAGCTGGCTGTGACACTTCCATGGAGACTGCCTGGCTGACAGCAAGACATGAT 2548

QY 97 GlyAspGlyLeuVal 101
Db 2549 GGAGATGGCTGGTG 2563

RESULT 4
ABQ82345
ID ABQ82345 standard; cDNA; 8473 BP.
XX AC ABQ82345;
XX DT 17-DEC-2002 (first entry)
XX DE Human NOV15c encoding cDNA SEQ ID NO:39.
XX KW Human; NOVX; cytotatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW anticancer; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 258..8144
XX FT /*tag= a
XX FT /product= "NOV15c"
XX PN WO200262999-A2.
XX PD 15-AUG-2002.
XX PF 31-DEC-2001; 2001WO-US049976.
XX PR 29-DEC-2000; 2000US-0258928P.
XX PR 02-JAN-2001; 2001US-0259415P.
XX PR 04-JAN-2001; 2001US-0259783P.
XX PR 20-FEB-2001; 2001US-0269814P.
XX PR 09-MAR-2001; 2001US-0279863P.
XX PR 29-MAR-2001; 2001US-02798332P.
XX PR 13-APR-2001; 2001US-0283689P.
XX PR 18-APR-2001; 2001US-0284447P.
XX PR 25-APR-2001; 2001US-0286683P.
XX PR 29-MAY-2001; 2001US-0294080P.
XX PR 16-AUG-2001; 2001US-0312915P.
XX PR 17-AUG-2001; 2001US-0313325P.
```

```
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX FA (CURA-) CURAGEN CORP.
XX FI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyanar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli RA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR,
XX WPI; 2002-732706/79.
DR P-PSDB; ABP53588.
XX PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX PS Claim 8; Page 119-121; 444pp; English.
XX CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytotatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15c, which is
CC located on chromosome 4
XX SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;
```

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 3.22e-32 | Length: | 8473 |
| Score: | 515.00 | Matches: | 84 |
| Percent Similarity: | 85.05% | Conservative: | 7 |
| Best Local Similarity: | 78.50% | Mismatches: | 10 |
| Query Match: | 83.47% | Indels: | 6 |
| DB: | 6 | Gaps: | 1 |

US-10-029-020-14_COPY_750_850 (1-101) x ABQ82345 (1-8473)

```
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2082 TGGACGGGCCAGCCTGTAATCAGAGAGCCTGCCACCCCGCTGTGCCGAGCAGGGACC 2141

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2142 TGCAGAGATGGCAAGTGTGAATGCAGCGGCTGGATGGAGACACTGCACACTCGT 2201

QY 41 HisTyrLeuAspArgVal-----LysGluGlyCysProGlyLeu 54
Db 2202 CACTATTGGATAAGATAGTTAAAGACAAGATAGGATATAAGAGGGTGTCTCTGTCG 2261

QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
Db 2262 TGCACAGCATGGAGATGTACCTTGCACCAAAATGGCGGACATTGTGTGTGCCAGCCT 2321
```

```
QY 75 GlyTTPArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp 94
Db 2322 GGATGGAGAGGAGCAGCTGTGACCTAGCCATGGAGACTCTTTCACAGATAGCAGGAC 2381

QY 95 AsnAspGlyAspGlyLeuVal 101
Db 2382 AATGAGGGGATGACTCATT 2402

RESULT 5
ABQ82344
ID ABQ82344 standard; cDNA; 8645 BP.
XX AC ABQ82344;
XX 17-DEC-2002 (first entry)
XX Human NOV15b encoding cDNA SEQ ID NO:37.
XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; autoimmune disease; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 151..8316
FT /*tag= a
FT /*product= "NOV15b"
XX
XX WO200262999-A2.
XX
XX 15-AUG-2002.
XX
XX 31-DEC-2001; 2001WO-US049976.
XX
XX 29-DEC-2000; 2000US-02589328P.
XX 02-JAN-2001; 2001US-0259415P.
XX 04-JAN-2001; 2001US-0259785P.
XX 20-FEB-2001; 2001US-0269814P.
XX 09-MAR-2001; 2001US-0279863P.
XX 29-MAR-2001; 2001US-0279832P.
XX 13-APR-2001; 2001US-0279833P.
XX 18-APR-2001; 2001US-0283889P.
XX 25-APR-2001; 2001US-0284447P.
XX 29-MAY-2001; 2001US-0286683P.
XX 16-AUG-2001; 2001US-0294080P.
XX 17-AUG-2001; 2001US-0312915P.
XX 17-SEP-2001; 2001US-0313325P.
XX 26-NOV-2001; 2001US-0322699P.
XX 26-NOV-2001; 2001US-0333350P.
XX (CURA-) CURAGEN CORP.
XX
XX Spytek KA, Li L, Wolenc AR, Vernet CM, Eisen A, Liu X;
PI Malyankar U, Shinkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SU, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
XX WPI; 2002-732706/79.
XX P-PSDB; ABP53587.
XX
XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
```

diseases.

Claim 8; Page 114-117; 444pp; English.

The present invention describes novel human proteins designated NOVX, where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective, cytostatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular, antidiabetic, antiinflammatory, antirheumatic, antiallergic, virucide, immunosuppressive, antianaemic, antianaemic, antibacterial, fungicide, protozoacide and antihelminthic activities, and can be used in gene therapy. The NOVX proteins, nucleotides or antibodies can be used in the manufacture of a medicament for treating a syndrome associated with a human disease selected from NOVX-associated disorder, such as cancers (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle, ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke, ischaemic cerebrovascular disease, Alzheimer's disease or Pick's disease), disorders of vesicular transport (e.g. cystic fibrosis, diabetes mellitus, Grave's disease, or goitre), gastrointestinal disorders (e.g. ulcerative colitis, or gastric and duodenal disorders), autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic and protozoal infections. The NOVX proteins can be used as immunogens to produce antibodies and as vaccines. The NOVX nucleotide sequences may be used in chromosome mapping, identifying individuals from minute biological samples (tissue typing), and in forensic identification of a biological sample. The present sequence encodes human NOV15b, which is located on chromosome 4

SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 3,29e-32 | Length: | 8645 |
| Score: | 515.00 | Matches: | 84 |
| Percent Similarity: | 85.05% | Conservative: | 7 |
| Best Local Similarity: | 78.50% | Mismatches: | 10 |
| Query Match: | 83.47% | Indels: | 6 |
| DB: | 6 | Gaps: | 1 |

US-10-029-020-14_COPY_750_850 (1-101) x ABQ82344 (1-8645)

| | | | |
|----------|--|--|------|
| QY | 1 | TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr | 20 |
| Db | 2254 | TGACGGGCCCGACCTGTAAATCAGAGAGCTGCCACCCCGCTGTGCCGACGCGGACC | 2313 |
| QY | 21 | CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla | 40 |
| Db | 2314 | TGCAAGGATGGCAAGTGTGATGCAGCCAGGGCTGGAATGGAGACACTGCATATCGCT | 2373 |
| QY | 41 | HisTyrLeuAspArgValVal-----LysGluGlyCysProGlyLeu | 54 |
| Db | 2374 | CACATTTGGATAAGATAGTTAAAGACAGATAGGATATAAAGAGGGTGTCTCTGCTCTG | 2433 |
| QY | 55 | CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu | 74 |
| Db | 2434 | TGCAACAGCATGGAAAGATGATACCTGGACCAAAATGGCGGACATGTGTGCCAGCCT | 2493 |
| QY | 75 | GlyTTPArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp | 94 |
| Db | 2494 | GGATGGAGAGGAGCAGGCTGTGACCTAGCCATGGAGACTCTTTGCACAGATAGCAGGAC | 2553 |
| QY | 95 | AsnAspGlyAspGlyLeuVal 101 | |
| Db | 2554 | AATGAAGGGGATGGACTCATT 2574 | |
| RESULT 6 | | | |
| ABQ82343 | | | |
| ID | ABQ82343 standard; cDNA; 8675 BP. | | |
| XX | AC ABQ82343; | | |
| XX | 17-DEC-2002 (first entry) | | |
| XX | Human NOV15a encoding cDNA SEQ ID NO:35. | | |

XX Human, NOVX; cytotatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 151..8328
XX /*tag= a
XX /product= "NOV15a"
XX /transl_except= (pos:1249..1251,aa:Ser)
XX
XX WO200262999-A2.
XX
XX 15-AUG-2002.
XX
XX 31-DEC-2001; 2001WO-US049976.
XX
XX 29-DEC-2000; 2000US-0259928P.
XX 02-JAN-2001; 2001US-0259415P.
XX 04-JAN-2001; 2001US-0259785P.
XX 20-FEB-2001; 2001US-0269814P.
XX 09-MAR-2001; 2001US-0279863P.
XX 29-MAR-2001; 2001US-0279832P.
XX 29-MAR-2001; 2001US-0279833P.
XX 13-APR-2001; 2001US-0283889P.
XX 18-APR-2001; 2001US-0284447P.
XX 25-APR-2001; 2001US-0286683P.
XX 29-MAY-2001; 2001US-0294080P.
XX 16-AUG-2001; 2001US-0312915P.
XX 17-AUG-2001; 2001US-0313325P.
XX 17-SEP-2001; 2001US-0322699P.
XX 26-NOV-2001; 2001US-0333350P.
XX (CURA-) CURAGEN CORP.
XX
XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
XX WPI; 2002-732706/79.
XX P-PSDB; ABP53586.
XX
XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
XX Claim 8; Page 110-112; 444pp; English.
XX
XX The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytotatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's

CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15a, which is
CC located on chromosome 4
XX
XX Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;
XX
XX Alignment Scores:
XX Pred. No.: 3.3e-32 Length: 8675
XX Score: 515.00 Matches: 84
XX Percent Similarity: 85.05% Conservative: 7
XX Best Local Similarity: 78.50% Mismatches: 10
XX Query Match: 83.47% Indels: 6
XX Gaps: 1
XX
XX US-10-029-020-14_COPY_750_850 (1-101) x ABQ82343 (1-8675)
XX
XX QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
XX Db 2266 TGGACGGGCCCGCTGTAATCAGAGAGCTGTCACCCCGCTGTGGCGAGACCGGACC 2325
XX
XX QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
XX Db 2326 TGCAGAGGATGCAAGTGTGAATGACCGAGCGCTGGAATGGAGACCTGCTACTATCGCT 2385
XX
XX QY 41 HisTyrLeuAspArgValVal-----LysGluGlyCysProGlyLeu 54
XX Db 2386 CACTATTGGATAGATAGTAGTAAACACAGATAGGATATAAGAGGGTGTCTCGTCTG 2445
XX
XX QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
XX Db 2446 TGCACAGCAATGGAAGATGTACCTGGACCAAAATGGCGGACATGTGTGTGCCAGCCT 2505
XX
XX QY 75 GlyTyrArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp 94
XX Db 2506 GGATGGAGAGGAGGAGGCTGTGACCTAGCATGGAGACTCTTTGGACAGATAGCAGGAC 2565
XX
XX QY 95 AsnAspGlyAspGlyLeuVal 101
XX Db 2566 AATGAGGGGATGGACTCATT 2586
XX
XX RESULT 7
XX ABQ82346
XX ID ABQ82346 standard; cDNA; 8487 BP.
XX AC ABQ82346;
XX DT 17-DEC-2002 (first entry)
XX XX
XX DE Human NOV15d encoding cDNA SEQ ID NO:41.
XX
XX KW Human; NOVX; cytotatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX
XX Homo sapiens.
XX Key Location/Qualifiers
XX FH


```
PT prostate tissue.
XX
PS Claim 22; Page 394-397; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences
XX
SQ Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6,46e-28 Length: 12879
Score: 464.50 Matches: 71
Percent Similarity: 82.00% Conservative: 11
Best Local Similarity: 71.00% Mismatches: 17
Query Match: 75.28% Indels: 1
DB: Gaps: 1
US-10-029-020-14_COPY_750_850 (1-101) x ABK92230 (1-12879)
QY 1 TrpMetGlyAlaAaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2210 TGGGTAGGACCAACATGTGAGGACCGCTCTCTCATTTCTTCTACTGAGCATGCCCAA 2269
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2270 TGCAGAGATGGAATGTGAGTGTAGCCCTGTGATGGAGGGGACCACTGCACATTTGCT 2329
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2330 CACTACTTAGAT---GCTGTCCGAGATGCTGCCAGGCTCTGCTTTTGGAAATGACGCA 2386
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2387 TGTACCCCTGGATCAAAATGGTTGGCACTGTGTGTGTCAGGTGGGTTGGAGTGGGACGCG 2446
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2447 TGCAGATGTTGTCATGGAATATGCTTTTGGAGATTAATTTGGACATATGAGATGTTTA 2506
RESULT 9
AAK51828
ID AAK51828 standard; cDNA; 13202 BP.
XX
AC AAK51828;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 373.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
FN WO200157190-A2.
XX
PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 01-JUL-2000; 2000US-00620325.
PR 15-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR P-PSDB; AAM78695.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 1; Page 1414-1426; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6,63e-28 Length: 13202
Score: 464.50 Matches: 71
Percent Similarity: 82.00% Conservative: 11
Best Local Similarity: 71.00% Mismatches: 17
Query Match: 75.28% Indels: 1
DB: Gaps: 1
US-10-029-020-14_COPY_750_850 (1-101) x AAK51828 (1-13202)
QY 1 TrpMetGlyAlaAaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2536 TGGGTAGGACCAACATGTGAGGACCGCTCTCTCATTTCTTCTACTGAGCATGCCCAA 2595
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2596 TGCAGAGATGGAATGTGAGTGTAGCCCTGTGATGGAGGGGACCACTGCACATTTGCT 2655
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2656 CACTACTTAGAT---GCTGTCCGAGATGCTGCCAGGCTCTGCTTTTGGAAATGACGCA 2712
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2713 TGTACCCCTGGATCAAAATGGTTGGCACTGTGTGTGTCAGGTGGGTTGGAGTGGGACGCG 2772
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2773 TGCAGATGTTGTCATGGAATATGCTTTTGGAGATTAATTTGGACATATGAGATGTTTA 2832
RESULT 10
```


AAAS14085
ID AAS14085 standard; DNA; 9826 BP.
AC AAS14085;
XX
DT 18-DEC-2001 (first entry)
DE Human FCTR3b DNA sequence.
XX
XX Human; FCTR3; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
KW astrocytoma; congenital neonatal alloimmune thrombocytopaenia; infection;
KW neurological disorder; neurodegenerative disorders; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greenow type I; Corneal dystrophy-lattice type I;
KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW anti-allergic; antiasthmatic; antiinfertility; antiinflammatory;
KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3b; neurestin-like protein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1. .279
FT /*tag= b
FT CDS 280. .8481
FT /*tag= a
FT /product= "Human FCTR3b"
FT 3'UTR 8482. .9826
FT /*tag= c
FT WO200166747-A2.
XX
PD 13-SEP-2001.
XX
XX 05-MAR-2001; 2001WO-US0007160.
XX
PR 03-MAR-2000; 2000US-0186592P.
PR 03-MAR-2000; 2000US-0186718P.
PR 06-MAR-2000; 2000US-0187293P.
PR 06-MAR-2000; 2000US-0187294P.
PR 17-MAR-2000; 2000US-0190400P.
PR 07-APR-2000; 2000US-0196018P.
PR 03-JAN-2001; 2001US-0259548P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
XX WPI: 2001-596837/67.
DR P-PSDB; AAU08680.
XX
PT Novel polypeptides designated as FCTR3 polypeptides, useful in detection,
PT prevention and treatment of a broad range of pathological states.
XX
XX Claim 9; Page 33-35; 215pp; English.
XX
XX The invention relates to human FCTR3 polypeptides, FCTR1-FCTR7, and the
XX nucleic acids encoding them. These sequences are useful for the treatment
XX or prevention of numerous disorders including myelogenous leukaemia,
XX carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
XX alloimmune thrombocytopaenia, neurological disorders, neurodegenerative
XX disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
XX -Tooth neuropathy, demyelinating Gardner syndrome, familial
XX myelodysplastic syndrome, mental health conditions, immunological

CC disorders, allergy and infection, bronchial asthma, Avellino type
CC eosinophilia, lung diseases, reproductive disorders, infertility, male
CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
CC Corneal dystrophy-Greenow type I, Corneal dystrophy-lattice type I and
CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
CC FCTR3b, a neurestin-like protein
XX
SQ Sequence 9826 BP; 2567 A; 2643 C; 2479 G; 2137 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3.49e-27 Length: 9826
Score: 454.00 Matches: 71
Percent Similarity: 77.23% Conservative: 7
Best Local Similarity: 70.30% Mismatches: 23
Query Match: 73.58% Indels: 0
DB: 5 Gaps: 0
US-10-029-020-14_COPY_750_850 (1-101) x AAS14085 (1-9826)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2434 TGGACAGGCGCAGCGTGTGACCGCGTGTGCCACCCCGCTGCAFTGAGCAGCGGACC 2493
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2494 TGTAAGATGGCAATGTGAATGCCGAGAGGGCTGGAAATGGTGAACACTGCACCATTTGGT 2553
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2554 AGCGAAACCGCAGCGCACCGAAACAGATGGCTGCCCTGACTTGTGCAACGGTAACGGGAGA 2613
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGluLeuGlyTrpArgGlyAlaGly 80
Db 2614 TGCACACTGGTTCAGAACAGCTGGCAGTGTCTGCCACCGCTGGAGAGGCCCGCGGA 2673
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2674 TGCAAGTTGGCCATGGAAACTTCCTGTCTGTATACACAGGATATAGGGAGATGGCTG 2733
QY 101 Val 101
Db 2734 GTG 2736
RESULT 11
ID ADB32023 standard; cDNA; 9826 BP.
XX
XX ADB32023;
XX
XX 04-DEC-2003 (first entry)
XX Human FCTR3b cDNA.
XX
XX Human; gene; ss; FCTR3b; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal alloimmune thrombocytopaenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurologic disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW lattice corneal dystrophy.
XX
OS Homo sapiens.
XX

```
PN US2003087816-A1.
XX 08-MAY-2003.
XX 05-MAR-2001; 2001US-00800198.
XX 03-MAR-2000; 2000US-0186592P.
XX (VERM/) VERMET C.
XX (PERN/) FERNANDES E.
XX (SHIM/) SHIMKETS R.
XX (HERR/) HERRMANN J.
XX (MAJU/) MAJUMDER K.
XX (MACD/) MACDOUGALL J.
XX (MISH/) MISHRA V.
XX (MEZE/) MEZES P S.
XX (RAST/) RASTELLI L.
XX Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
PI MacDougall J, Mishra V, Mezes PS, Rastelli L;
XX P-PSDB; ADB32024.
XX WPI; 2003-625633/59.
XX DR P-PSDB; ADB32024.
XX New FCTR3f polypeptide and encoding polynucleotide, useful for preventing
XX or treating FCTR3f-related disorders, such as cancer, autoimmune,
XX neurodegenerative, gastrointestinal, reproductive and inflammatory
XX diseases.
XX Claim 9; Page 29-32; 155pp; English.
XX The invention relates to FCTR3f polypeptides and the polynucleotides
XX encoding them. The sequences of the invention are useful for the
XX manufacture of a medicament for diagnosing and treating disorders
XX associated with the FCTR3f polypeptide, such as colorectal cancer,
XX adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
XX autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
XX tumours, mammary tumours, human gliomas, astrocytomas, renal cell
XX carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
XX cell and granular cell carcinomas, neurological disorders,
XX neurodegenerative disorders, nerve trauma, familial myelodysplastic
XX syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
XX conditions, immunological disorders, allergy and infection, asthma, lung
XX diseases, male and female reproductive disorders, deafness, glycoprotein
XX deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
XX C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
XX infection, spinocerebellar ataxia, plasmodium falciparum infection,
XX Greenough's corneal dystrophy and lattice corneal dystrophy. This sequence
XX represents cDNA encoding an FCTR3f polypeptide of the invention.
XX SQ Sequence 9826 BP; 2567 A; 2645 C; 2476 G; 2138 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3,49e-27 Length: 9826
Score: 454.00 Matches: 71
Percent Similarity: 77.23% Conservative: 7
Best Local Similarity: 70.30% Mismatches: 23
Query Match: 73.58% Indels: 0
DB: Gaps: 0
US-10-029-020-14_COPY_750_850 (1-101) x ADB32023 (1-9826)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
DB 2434 TGGACAGGCGCAGCGGTGACACAGCGCGTGTGCACCCCGCGCTGTCAGTGGACGAGC 2493
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
DB 2494 TGTAAAGATGGCAATGTGATGCCGAGAGGCGTGGATGTGAACACTGCACCAITGGT 2553
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyArg 60
DB 2554 AGGCAAAACGCGAGCGCCAGCAACAGATGGTGCCTGACTTGTGCAACGGTAAACGGAGA 2613
```

```
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyValAlaGly 80
DB 2614 TGCACACTGGGTGAGAACAGCTGGCAGCTGTCTGCGCCAGACCGGCTGGAGAGGCGCCGGA 2673
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
DB 2674 TGCAACGTTGCCATGGAAACTTCTGTGCTGATAACAAGGATAATGAGGGAGATGGCCTG 2733
QY 101 Val 101
DB 2734 GTG 2736
RESULT 12
AAS14089
ID AAS14089 standard; DNA; 9729 BP.
AC AAS14089;
XX 18-DEC-2001 (first entry)
XX Human FCTR3f DNA sequence.
XX Human; FCTR3f; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
XX astrocytoma; congenital neonatal autoimmune thrombocytopenia; infection;
XX neurological disorder; neurodegenerative disorders; nerve trauma;
XX familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
XX demyelinating Gardner syndrome; familial myelodysplastic syndrome;
XX mental health condition; immunological disorder; allergy; infertility;
XX bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
XX reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
XX desmoid disease; pancreatic disease; Schistosoma mansoni infection;
XX gastric disorders; plasmodium falciparum parasitaemia; diabetes;
XX Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; type I;
XX Corneal dystrophy-Greenough type I; Corneal dystrophy-lattice type I;
XX Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
XX anti-allergic; antiasthmatic; antifertility; antinflammatory;
XX antidiabetic; protooncogene; hepatotropic; virucide; ophthalmological;
XX gynaecological; antifertility; immunostimulant; auditory; haemostatic;
XX gene therapy; FCTR3a; neurestin-like protein; FCTR3f.
XX Homo sapiens.
XX Key Location/Qualifiers
XX 5'UTR 1..209
XX /tag= b
XX CDS 210..8384
XX /tag= a
XX /product= "Human FCTR3f"
XX 3'UTR 8385..9729
XX /tag= c
XX WO200166747-A2.
XX 13-SEP-2001.
XX 05-MAR-2001; 2001WO-US007160.
XX 03-MAR-2000; 2000US-0186592P.
XX 03-MAR-2000; 2000US-0186718P.
XX 06-MAR-2000; 2000US-0187293P.
XX 06-MAR-2000; 2000US-0187294P.
XX 17-MAR-2000; 2000US-0190400P.
XX 07-APR-2000; 2000US-0196018P.
XX 03-JAN-2001; 2001US-0259548P.
XX (CURA-) CURAGEN CORP.
XX Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
PI MacDougall J, Mishra V, Mezes PS, Rastelli L;
XX WPI; 2001-596837/67.
XX P-PSDB; AAU08681.
```

XX Novel polypeptides designated as FCTR3f polypeptides, useful in detection,
PT prevention and treatment of a broad range of pathological states.
XX Claim 9; Page 37-39; 215pp; English.
XX The invention relates to human FCTR3 polypeptides, FCTR1-FCTR7, and the
CC nucleic acids encoding them. These sequences are useful for the treatment
CC or prevention of numerous disorders including myelogenous leukaemia,
CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
CC autoimmune thrombocytopenia, neurological disorders, neurodegenerative
CC -Tooth neuropathy, demyelinating Gardner syndrome, Charcot-Marie
CC myelodysplastic syndrome, mental health conditions, immunological
CC disorders, allergy and infection, bronchial asthma, Avellino type
CC eosinophilia, lung diseases, reproductive disorders, infertility, male
CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC infection, Spino cerebellar ataxia, Plasmodium falciparum parasitaemia,
CC Corneal dystrophy-Greenou type I, Corneal dystrophy-lattice type I and
CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
CC the FCTR3a homologue FCTR3f
XX SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.7e-27 Length: 9729
Score: 448.50 Matches: 71
Percent Similarity: 77.23% Conservative: 7
Best Local Similarity: 70.30% Mismatches: 14
Query Match: 72.69% Indels: 9
DB: 5 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x AAS14089 (1-9729)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2364 TGGACAGGCGCAGCGTGTGACACGCGGTGTGCACCCCGCTGCATTGAGCATGGAGCC 2423
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2424 TGTAAAGATGGCAATGTGAATGCCGAGAGGGCTGGAATGTCACACTGCACCAT--- 2480
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2481 -----GATGGCTGCCCTGACTTTGTCAACGGTAACGGGAGA 2516
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2517 TGCACACTGGGTGAGAACAGCTGGCAGTGTGTCTGCCAGACCGCTGGAGAGGCCCGGA 2576
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2577 TGCACAGTTGCCATGGAAATCTCTGTGTGATACAGGATAAATGAGGAGATGGCCTG 2636
QY 101 Val 101
Db 2637 GTG 2639

RESULT 13

ADB32028

ID ADB32028 standard; cDNA; 9729 BP.

AC ADB32028;

DT 04-DEC-2003 (first entry)

DE Human FCTR3f cDNA.

XX

XX Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;

KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Greenou's corneal dystrophy;
XX lattice corneal dystrophy.

XX Homo sapiens.

XX US2003087816-A1.

XX 08-MAY-2003.

XX 05-MAR-2001; 2001US-00800198.

XX 03-MAR-2000; 2000US-0186592P.

XX (VERM/) VERMET C.

XX (PERN/) FERNANDES E.

XX (SHIM/) SHIMKETS R.

XX (HERR/) HERRMANN J.

XX (MAJU/) MAJUMDER K.

XX (MACD/) MACDOUGALL J.

XX (MISH/) MISHRA V.

XX (MEZE/) MEZES P S.

XX (RAST/) RASTELLI L.

XX Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
XX MacDougall J, Mishra V, Mezes PS, Rastelli L;

XX WPI: 2003-625633/59.

XX P-PSDB; ADB32029.

XX New FCTR3 polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTR3-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.

XX Claim 9; Page 34-37; 155pp; English.

XX The invention relates to FCTR3 polypeptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTR3 polypeptide, such as colorectal cancer,
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, neuroendocrine disorders,
CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
CC conditions, immunological disorders, allergy and infection, asthma, lung
CC diseases, male and female reproductive disorders, deafness, glycoprotein
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC Greenou's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTR3 polypeptide of the invention.

XX SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.7e-27 Length: 9729
Score: 448.50 Matches: 71
Percent Similarity: 77.23% Conservative: 7
Best Local Similarity: 70.30% Mismatches: 14
Query Match: 72.69% Indels: 9
DB: 5 Gaps: 1

| | |
|---|---|
| US-10-029-020-14_COPY_750_850 (1-101) x ADB32028 (1-9729) | |
| QY | 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20 |
| Db | 2364 TGGACAGCGCAGCGTGTGACACAGCGCTGTGCCACCCCGCTGCAITGAGCATGGGACC 2423 |
| QY | 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40 |
| Db | 2424 TGTAAGATGGCAATGTGAATGCCGAGAGCGCTGGAATGTGAACACTGCACCATTT--- 2480 |
| QY | 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60 |
| Db | 2481 -----GATGGCTGCCCTGACTTGTGCAACGGTAACGGGAGA 2516 |
| QY | 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80 |
| Db | 2517 TGCACACTGGCTCAGACAGCTGGCAGTGTGTCTGCCAGACCGGCTGCAGAGGGCCCGGA 2576 |
| QY | 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100 |
| Db | 2577 TGCACAGTTGCCATGAACTTCTGTCTGATACAAAGGATAATGAGGAGATGGCCTG 2636 |
| QY | 101 Val 101 |
| Db | 2637 GTG 2639 |
| RESULT 14 | |
| AAL60066 | |
| ID | AAL60066 standard; cDNA; 4245 BP. |
| XX | |
| AC | AAL60066; |
| DT | 27-AUG-2003 (first entry) |
| DE | Human PC099 cDNA. |
| XX | |
| KW | Human; differentially regulated protein; prevention; therapy; vaccine; |
| KW | prostate cancer; gene therapy; PC099; teneurin-2; chromosome 5q34; gene; |
| KW | ss. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| Key | Location/Qualifiers |
| FT | 1..4056 |
| FT | /*tag= a |
| FT | /product= "Human PC099 protein" |
| FT | /note= "No start codon" |
| FT | /partial |
| XX | |
| PN | WO2003040331-A2. |
| XX | |
| PD | 15-MAY-2003. |
| XX | |
| PF | 07-NOV-2002; 2002WO-US035563. |
| XX | |
| PR | 07-NOV-2001; 2001US-0331041P. |
| PR | 07-NOV-2001; 2001US-0331042P. |
| PR | 18-DEC-2001; 2001US-0340251P. |
| PR | 07-JAN-2002; 2002US-0344791P. |
| XX | |
| PA | (ORIG-) ORIGENE TECHNOLOGIES INC. |
| XX | |
| PI | Sun Z, Li X, Jay G, Kovacs KF, Fan W; |
| XX | |
| WI | PI: 2003-449451/42. |
| DR | P-PSDB; AA029571. |
| XX | |
| XX | New polynucleotide for diagnosing, staging, monitoring, prognosticating, |
| PT | preventing or treating, or determining the predisposition to, diseases or |
| PT | conditions such as prostate cancer, and for research or forensic science. |
| XX | |
| XX | Claim 29; Page 158-163; 100pp; English. |

| | | |
|---|---|------------------|
| XX | The present invention relates to novel differentially regulated genes and | |
| CC | polypeptides encoded by them. Sequences of the invention are useful in | |
| CC | diagnosing, staging, monitoring, prognosticating, preventing, treating or | |
| CC | determining the predisposition to diseases or conditions such as prostate | |
| CC | cancer. They may be used as molecular markers, drug targets, vaccines, in | |
| CC | gene therapy, research, clinical medicine or forensic science. The | |
| CC | present sequence is a differentially regulated prostate cDNA, PC099 which | |
| CC | codes for teneurin-2. PC099 gene is located on chromosome 5q34 | |
| XX | | |
| SQ | Sequence 4245 BP; 1035 A; 1132 C; 1115 G; 963 T; 0 U; 0 Other; | |
| Alignment Scores: | | |
| Pred. No.: | 1.66e-17 | Length: 4245 |
| Score: | 330.50 | Matches: 60 |
| Percent Similarity: | 57.36% | Conservative: 14 |
| Best Local Similarity: | 46.51% | Mismatches: 26 |
| Query Match: | 53.57% | Indels: 29 |
| DB: | 7 | Gaps: 4 |
| US-10-029-020-14_COPY_750_850 (1-101) x AAL60066 (1-4245) | | |
| QY | 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCys---HisProArgCysAlaGluHisGly 19 | |
| Db | 790 TACAAAGCGGAGCACTGTGAGGAAGTTGATTGCTTGGATCCCACTGCTCCAGCCAGGA 849 | |
| QY | 20 ThrCysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle 39 | |
| Db | 850 GTCTGTGTGAATGGAGAATGCTGTGCAGCCCTGGCTGGGTGTCTGAACCTGTGAGCTG 909 | |
| QY | 40 AlaHis-----TyrLeu----- 43 | |
| Db | 910 GCGAGGTTCCAGTGCCAGACCGACGTGCGGATGCGACGTACTGCTGCACACGGGC 969 | |
| QY | 44 -----AspArgValValLysGluGlyCysPro 52 | |
| Db | 970 CTCTCAGCTCGCATCCCACTGGATGGGTCCCGACTGCTCTGTT---GATGGTGCCTC 1026 | |
| QY | 53 GlyLeuCysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCys 72 | |
| Db | 1027 GACTTGTGCACCGTAACGGTAACGGAGATGCACACTGGGTGACAGACAGCTGGCAGTGTCTGC 1086 | |
| QY | 73 GlnLeuGlyTrpArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSer 92 | |
| Db | 1087 CAGACCGCTGGAGAGGGCCCGATGCAACAGTTGCCATGGAACCTCTCTGTGCTGATAAC 1146 | |
| QY | 93 LysAspAsnAspGlyAspGlyLeuVal 101 | |
| Db | 1147 AAGGATAATGAGGGAGATGGCCTGGTG 1173 | |
| RESULT 15 | | |
| ACC72051 | | |
| ID | ACC72051 standard; DNA; 9058 BP. | |
| XX | | |
| AC | ACC72051; | |
| XX | | |
| DT | 08-JUL-2003 (first entry) | |
| XX | | |
| DE | BCU0205A gene #SEQ ID 79. | |
| XX | | |
| KW | Breast cancer; cytostatic; gene therapy; antisense therapy; regulated; | |
| KW | drug discovery; clinical medicine; forensic medicine; gene; | |
| KW | chromosome 5q33.3; ds. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | WO2003029421-A2. | |
| XX | | |
| PD | 10-APR-2003. | |
| XX | | |
| PF | 02-OCT-2002; 2002WO-US031287. | |
| XX | | |
| PR | 03-OCT-2001; 2001US-0326526P. | |

PR 14-MAY-2002; 2002US-00144194.
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
XX PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
XX DR WPI; 2003-381623/36.
XX DR P-PsDB; ABR58317.
XX
XX PT New isolated human differentially-regulated breast cancer polynucleotide
XX PT and polypeptide, useful for diagnosing, staging, prognosticating,
XX PT preventing and/or treating diseases and conditions relating to breast
XX PT cancer.
XX
XX PS Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.
XX
XX CC The invention relates to isolated polynucleotides which are
XX CC differentially-regulated in breast cancer. The methods and compositions
XX CC of the present invention are useful for detecting, diagnosing, staging,
XX CC monitoring, prognosticating, preventing and/or treating diseases and
XX CC conditions relating to breast cancer, and may be used in gene therapy or
XX CC antisense therapy. They can also be used in research, drug discovery,
XX CC clinical medicine and forensic medicine. Sequences given in records
XX CC ACC72012-ACC72074 represent polynucleotides of the invention that are
XX CC differentially-regulated in breast cancer. NOTE: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.67e-17 Length: 9058
Score: 330.50 Matches: 60
Percent Similarity: 57.36% Conservative: 14
Best Local Similarity: 46.51% Mismatches: 26
Query Match: 53.57% Indels: 29
DB: 7 Gaps: 4

US-10-029-020-14_COPY_750_850 (1-101) x ACC72051 (1-9058)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCys---HisProArgCysAlaGluHisGly 19
Db 1571 TACAAAGGCGAGCAGCTGTGAGGAATGATTGCTTGGATCCACCTGCTCCAGCCACGGA 1630
QY 20 ThrCysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle 39
Db 1631 GTCTGTGTGATGAGATGCCCTGTGCAGCCCTGGCTGGTCTGAACCTGTGAGCTG 1690
QY 40 AlaHis-----TyrLeu----- 43
Db 1691 GCGAGGGTCCAGTCCAGACCAGTCAGTGGGCATGGCAGCTACTGCTGACACGGGC 1750
QY 44 -----AspArgValValLysGluGlyCysPro 52
Db 1751 CTCTGCAGTCCAGTCCCACTGGATGGGTCCCGACTGCTCTGTT---GATGGCTGCCCT 1807
QY 53 GlyLeuCysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCys 72
Db 1808 GACTTGTGCAACGTTNACGGGAGATGCACACTGGGTGCAGAACAGCTGGCAGTGTGTCTGC 1867
QY 73 GlnLeuGlyTrpArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSer 92
Db 1868 CAGACCGGTGGAGAGGGCCCGATGCACTGGTGCATGGAACTTCCTGTGCTGATAAC 1927
QY 93 LysAspAsnAspGlyAspGlyLeuVal 101
Db 1928 AAGGATAATGAGGAGATGCCCTGGTG 1954

Search completed: August 14, 2004, 02:29:24
Job time : 200.323 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 142.298 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_450_520

Perfect score: 372

Sequence: 1 TFWRSQVFIDHPVHLKFVNS.....RLITQEARSLGTFGRQSRGT 71

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DRV=xlh

-Q=/cgn2_1/USPro.spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519

-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62

-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100

-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0

-MAXLEN=2000000000 -USER=US10029020@cgn_1_2156_@runat_06082004_112218_29331

-NCFU=6 -ICPU=3 -NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100

-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
|------------|-------|-------------|--------|----|-------------|

ALIGNMENTS

RESULT 1

US-10-383-201-45

; Sequence 45, Application US/10383201

; Publication No. US2004002926A1

; GENERAL INFORMATION:

; APPLICANT: Alsobrook II, John et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS

; FILE REFERENCE: 21402-568A

; CURRENT APPLICATION NUMBER: US/10/383,201

; CURRENT FILING DATE: 2003-03-06

; PRIOR APPLICATION NUMBER: 10/029020

; PRIOR FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/365,984

; PRIOR FILING DATE: 2002-03-20

; PRIOR APPLICATION NUMBER: 60/372,022

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/389,143

; PRIOR FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 60/391,779

; PRIOR FILING DATE: 2002-06-26

; PRIOR APPLICATION NUMBER: 60/410,755

; PRIOR FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: 60/412,957

| | | | | | | |
|----|-------|-------|-------|----|-------------------|-------------------|
| 1 | 372 | 100.0 | 1392 | 13 | US-10-383-201-45 | Sequence 45, Appl |
| 2 | 372 | 100.0 | 1392 | 13 | US-10-383-201-53 | Sequence 53, Appl |
| 3 | 372 | 100.0 | 1476 | 13 | US-10-383-201-41 | Sequence 41, Appl |
| 4 | 372 | 100.0 | 1680 | 15 | US-10-290-578-3 | Sequence 3, Appl |
| 5 | 372 | 100.0 | 2387 | 15 | US-10-290-578-11 | Sequence 11, Appl |
| 6 | 372 | 100.0 | 3111 | 9 | US-09-773-517-12 | Sequence 12, Appl |
| 7 | 372 | 100.0 | 3111 | 9 | US-09-792-025-12 | Sequence 12, Appl |
| 8 | 372 | 100.0 | 3111 | 9 | US-09-849-868-12 | Sequence 12, Appl |
| 9 | 372 | 100.0 | 3111 | 15 | US-10-290-578-1 | Sequence 1, Appl |
| 10 | 372 | 100.0 | 3111 | 15 | US-10-453-183-12 | Sequence 12, Appl |
| 11 | 372 | 100.0 | 8354 | 13 | US-10-383-201-43 | Sequence 43, Appl |
| 12 | 372 | 100.0 | 8354 | 13 | US-10-029-020-13 | Sequence 13, Appl |
| 13 | 372 | 100.0 | 8355 | 13 | US-10-383-201-55 | Sequence 55, Appl |
| 14 | 372 | 100.0 | 8438 | 13 | US-10-042-865-1 | Sequence 1, Appl |
| 15 | 254 | 68.3 | 8473 | 17 | US-10-038-854-39 | Sequence 39, Appl |
| 16 | 254 | 68.3 | 8487 | 17 | US-10-038-854-41 | Sequence 41, Appl |
| 17 | 254 | 68.3 | 8645 | 17 | US-10-038-854-37 | Sequence 37, Appl |
| 18 | 234 | 62.9 | 8675 | 17 | US-10-038-854-35 | Sequence 35, Appl |
| 19 | 220 | 59.1 | 12880 | 16 | US-10-295-027-927 | Sequence 927, App |
| 20 | 194.5 | 52.3 | 1430 | 9 | US-09-808-602-5 | Sequence 5, Appl |
| 21 | 194.5 | 52.3 | 1430 | 10 | US-09-800-198-5 | Sequence 5, Appl |
| 22 | 194.5 | 52.3 | 1431 | 10 | US-09-977-418-21 | Sequence 21, Appl |
| 23 | 194.5 | 52.3 | 1431 | 10 | US-09-977-033A-21 | Sequence 21, Appl |
| 24 | 194.5 | 52.3 | 1431 | 10 | US-09-977-753C-21 | Sequence 21, Appl |
| 25 | 194.5 | 52.3 | 1431 | 10 | US-09-977-639A-21 | Sequence 21, Appl |
| 26 | 194.5 | 52.3 | 1431 | 11 | US-09-977-819B-21 | Sequence 21, Appl |
| 27 | 194.5 | 52.3 | 1727 | 16 | US-10-094-749-710 | Sequence 710, App |
| 28 | 194.5 | 52.3 | 1743 | 15 | US-10-098-871-7 | Sequence 7, Appl |
| 29 | 194.5 | 52.3 | 2496 | 9 | US-09-808-602-75 | Sequence 75, Appl |
| 30 | 194.5 | 52.3 | 2496 | 10 | US-09-800-198-63 | Sequence 63, Appl |
| 31 | 194.5 | 52.3 | 8409 | 9 | US-09-808-602-79 | Sequence 79, Appl |
| 32 | 194.5 | 52.3 | 8409 | 10 | US-09-800-198-67 | Sequence 67, Appl |
| 33 | 194.5 | 52.3 | 8575 | 13 | US-10-072-012-143 | Sequence 143, App |
| 34 | 194.5 | 52.3 | 8689 | 9 | US-09-808-602-78 | Sequence 78, Appl |
| 35 | 194.5 | 52.3 | 8689 | 10 | US-09-800-198-66 | Sequence 66, Appl |
| 36 | 194.5 | 52.3 | 8797 | 9 | US-09-808-602-74 | Sequence 74, Appl |
| 37 | 194.5 | 52.3 | 8797 | 9 | US-09-808-602-77 | Sequence 77, Appl |
| 38 | 194.5 | 52.3 | 8797 | 10 | US-09-800-198-62 | Sequence 62, Appl |
| 39 | 194.5 | 52.3 | 8797 | 10 | US-09-800-198-65 | Sequence 65, Appl |
| 40 | 194.5 | 52.3 | 9058 | 16 | US-10-144-194A-79 | Sequence 79, Appl |
| 41 | 194.5 | 52.3 | 9695 | 16 | US-10-144-194A-81 | Sequence 81, Appl |
| 42 | 194.5 | 52.3 | 9729 | 9 | US-09-808-602-12 | Sequence 12, Appl |
| 43 | 194.5 | 52.3 | 9729 | 10 | US-09-800-198-12 | Sequence 12, Appl |
| 44 | 194.5 | 52.3 | 9826 | 9 | US-09-808-602-7 | Sequence 7, Appl |
| 45 | 194.5 | 52.3 | 9826 | 10 | US-09-800-198-7 | Sequence 7, Appl |

```
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 45
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1392)
US-10-383-201-45
Alignment Scores:
Pred. No.: 1,01e-47 Length: 1392
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x US-10-383-201-45 (1-1392)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
DB 244 ACTTTCGGAGATCTCAAGTGTTCATAGACCATCTCTGTCATCTGAAATTCATGTGCT 303
QY 21 LeuGlyLysAlaAlaLeuValGlyTleTyrGlyArgLysGlyLeuProProSerHisThr 40
DB 304 CTGGAAAGGAGCCCTGGTGGCATTATGTCAGAAAAGGCTCCCTCCCTTCACATACA 363
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
DB 364 CAGTTTGACTTTGTGGAGCTGCTGGATGCGAGAGGCTCTTAACCCAGAGGCGCGGAGC 423
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
DB 424 CTAGAGGGGAGCCCGCGCCAGTCTCGGGGAAC 456
RESULT 2
US-10-383-201-53
; Sequence 53, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 41
; LENGTH: 1476
; TYPE: DNA
```

```
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 45
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1392)
US-10-383-201-45
Alignment Scores:
Pred. No.: 1,01e-47 Length: 1392
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x US-10-383-201-45 (1-1392)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
DB 244 ACTTTCGGAGATCTCAAGTGTTCATAGACCATCTCTGTCATCTGAAATTCATGTGCT 303
QY 21 LeuGlyLysAlaAlaLeuValGlyTleTyrGlyArgLysGlyLeuProProSerHisThr 40
DB 304 CTGGAAAGGAGCCCTGGTGGCATTATGTCAGAAAAGGCTCCCTCCCTTCACATACA 363
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
DB 364 CAGTTTGACTTTGTGGAGCTGCTGGATGCGAGAGGCTCTTAACCCAGAGGCGCGGAGC 423
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
DB 424 CTAGAGGGGAGCCCGCGCCAGTCTCGGGGAAC 456
RESULT 2
US-10-383-201-53
; Sequence 53, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 41
; LENGTH: 1476
; TYPE: DNA
```



```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1464)
US-10-383-201-41
Alignment Scores:
Pred. No.: 11e-47 Length: 1476
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x US-10-383-201-41 (1-1476)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 322 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCATCTGAATTCATATGTGTCT 381
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 382 CTGGGAAGGAGCGCCCTGGTGGCATTTATGGCAAGAAAGCCCTCCCTCTCATACATA 441
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 442 CAGTTTCACTTTGTGGAGCTCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAGC 501
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 502 CTAGAGGGGACCCCGCCAGCTCTCGGGGAAC 534
RESULT 4
US-10-290-578-3
; Sequence 3, Application US/10290578
; Publication No. US20030078389A1
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/290,578
; FILING DATE: 08-Jul-96
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/021640
; FILING DATE: 12-Jul-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1680 base pairs
```

```
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-290-578-3
Alignment Scores:
Pred. No.: 1.3e-47 Length: 1680
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x US-10-290-578-3 (1-1680)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1348 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCATCTGAATTCATATGTGTCT 1407
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1408 CTGGGAAGGAGCGCCCTGGTGGCATTTATGGCAAGAAAGCCCTCCCTCTCATACATA 1467
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 1468 CAGTTTCACTTTGTGGAGCTCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAGC 1527
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1528 CTAGAGGGGACCCCGCCAGCTCTCGGGGAAC 1560
RESULT 5
US-10-290-578-11
; Sequence 11, Application US/10290578
; Publication No. US20030078389A1
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/290,578
; FILING DATE: 08-Jul-96
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/021640
; FILING DATE: 12-Jul-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
```

```

Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1893

RESULT 7
US-09-792-025-12
; Sequence 12, Application US/09792025
; Patent No. US20020042087A1
; GENERAL INFORMATION:
; APPLICANT: Sliwkowski, Mark
; Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09792,025
; FILING DATE: 23-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/020,598
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwartz, Timothy R.
; REGISTRATION NUMBER: 32171
; REFERENCE/DOCKET NUMBER: P1145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-7467
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: nucleic acid
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-792-025-12

Alignment Scores:
Pred. No.: 2,96e-47 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-09-792-025-12 (1-3111)
Qy 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuIysPhe 71
Db 1681 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCTCTGTGCATCTGAAAT 1741
Qy 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgIysGlyLeuProPhe 71
Db 1741 CTGGGAAAGCGACGCCCTGGTTGGCATTTATGGCAGAAAGCGCTCCCTCC 1791
Qy 41 GlnPheAspPheValGlnLeuLeuAspGlyArgGlnLeuThrGlnG 71
Db 1801 CAGTTTGGACTTTGTGAGCTGCTGGATGGCAGGAGGCTCTTACCCAG 1851
Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71

```

```
Db      1861 CTAGAGGGGACCCGCGCCAGTCTCGGGAACT 1893
|||||
RESULT 8
US-09-849-868-12
; Sequence 12, Application US/09849868
; Patent No. US20020081299A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Gao, Wei-Qiang
; TITLE OF INVENTION: HAIR CELL DISORDERS
; FILE REFERENCE: GENENT.035VPC
; CURRENT APPLICATION NUMBER: US/09/849,868
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (334)...(2637)
US-09-849-868-12

Alignment Scores:
Pred. No.:      2,96e-47      Length:      3111
Score:          372.00      Matches:      71
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              9          Gaps:          0

US-10-029-020-14_COPY_450_520 (1-71) x US-09-849-868-12 (1-3111)
QY      1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db      1681 ACTTTCGGAGATCTCAAGTGTTTCATAGACCATCTCGTGCATCTGAATTCATGTGTCT 1740
QY      21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db      1741 CTGGGAAGGACGACCCCTGGTGGCATTTATGGCAGAAAGGCCCTCCCTCCATACATACA 1800
QY      41 GlnPheAspPheValGluLeuLeuAspGlyArgGlnThrGlnGluAlaArgSer 60
Db      1801 CAGTTTGACTTTGTGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGGCGCGAGC 1860
QY      61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db      1861 CTAGAGGGGACCCGCGCCAGTCTCGGGAACT 1893

RESULT 9
US-10-290-578-1
; Sequence 1, Application US/10290578
; Publication No. US20030078389A1
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/10/290,578
; FILING DATE: 08-No. US20030078389A1-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/021640
; FILING DATE: 12-Jul-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: PL043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-290-578-1

Alignment Scores:
Pred. No.:      2,96e-47      Length:      3111
Score:          372.00      Matches:      71
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              15         Gaps:          0

US-10-029-020-14_COPY_450_520 (1-71) x US-10-290-578-1 (1-3111)
QY      1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db      1681 ACTTTCGGAGATCTCAAGTGTTTCATAGACCATCTCGTGCATCTGAATTCATGTGTCT 1740
QY      21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db      1741 CTGGGAAGGACGACCCCTGGTGGCATTTATGGCAGAAAGGCCCTCCCTCCATACATACA 1800
QY      41 GlnPheAspPheValGluLeuLeuAspGlyArgGlnThrGlnGluAlaArgSer 60
Db      1801 CAGTTTGACTTTGTGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGGCGCGAGC 1860
QY      61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db      1861 CTAGAGGGGACCCGCGCCAGTCTCGGGAACT 1893

RESULT 10
US-10-453-183-12
; Sequence 12, Application US/10453183
; Publication No. US20030199429A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; FILE REFERENCE: P1145R1
; CURRENT APPLICATION NUMBER: US/10/453,183
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: US/09/243,198
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,866
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-453-183-12
```

```
Alignment Scores:
Pred. No.: 2,96e-47 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-10-453-183-12 (1-3111)

Qy 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTTCGGAGATCTCAAGTGTTCATAGACCATCTGTGTCATCTGAAATTCAAATGTGTCT 1740

Qy 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1741 CTGGAAAGGAGCCCTGGTGGCATTTATGGCAGAAAAGGCTCCCTTCACATACA 1800

Qy 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTGTGGAGCTGCTGATGCGCAGGAGCTCTTAACCCAGAGGCGCGGAGC 1860

Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCGCAGTCTCGGGGAAC 1893

RESULT 11
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 1,11e-46 Length: 8354
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-10-029-020-13 (1-8354)

Qy 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1382 ACTTTCGGAGATCTCAAGTGTTCATAGACCATCTGTGTCATCTGAATTCAAATGTGTCT 1441

Qy 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1442 CTGGAAAGGAGCCCTGGTGGCATTTATGGCAGAAAAGGCTCCCTTCACATACA 1501

Alignment Scores:
Pred. No.: 1,11e-46 Length: 8354
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-10-029-020-13 (1-8354)

; Sequence 13, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-029-020-13
```

```

Qy 41 GlnPheAspPheValGlnLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1502 CAGTTTGACTTTCTGGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAGC 1561
Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1562 CTAGAGGGACCCCGCGCCAGTCTCGGGGAAC 1594

RESULT 13
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS OF USE
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:
Pred. No.: 1,11e-46 Length: 8355
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-10-383-201-55 (1-8355)

Qy 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1348 ACTTTCTGGAGATCTCAAGTGTTCTATAGACCATCTCTGTGCATCTGAAATTCATGTGCT 1407
Qy 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 1408 CTGGGAAGGAGCGCCCTGGTTGGCATTTATGGCAGAAAAGGCTCCCTCTCACATACA 1467
Qy 41 GlnPheAspPheValGlnLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1468 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAGC 1527
Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71

```

[illegible]

| QY | 1 | ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuIysPheAsnValSer | 20 |
|----|------|--|------|
| Db | 1358 | ACTTTCGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT | 1417 |

```
Qy 21 LeuGlyLysAlaLeuValGlyIleTyGlyArgLysGlyLeuProProSerHisThr 40
Db 1418 CTGGGAAAGCAGCCCTGGTGGCATTTATGGCAGAAAAGGCCCTCCCTTCACATACA 1477

Qy 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 1478 CAGTTTGACTTTGGAGCTCTGGATGGCAGGAGGCTCCTTAACCCAGGAGCGCGGAGC 1537

Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1538 CTAGAGGGGACCCCGCGCCAGCTCTCGGGGAAC 1570

RESULT 15
US-10-038-854-39
; Sequence 39, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shinkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39

; LENGTH: 8473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-39

Alignment Scores:
Pred. No.: 3,1e-28 Length: 8473
Score: 254.00 Matches: 48
Percent Similarity: 80.88% Conservative: 7
Best Local Similarity: 70.59% Mismatches: 13
Query Match: 68.28% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-10-038-854-39 (1-8473)
Qy 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 1185 TTCTGGAGATCACAGCTCTTCATTGATCAGCCAGAGTTTCTTAATTCATCTCTT 1244
Qy 22 GlyLysAlaLeuValGlyIleTyGlyArgLysGlyLeuProProSerHisThrGln 41
Db 1245 CAGAGGATGATGATTGGAGTATATGCCGGAAGGCTTACCGCTTCCCATACTCAG 1304
Qy 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSerLeu 61
Db 1305 TATGACTTCGTGGAGCTCTCGATGGCAGCAGGCTGATTGCCAGAGAGCAGCGGAGCTG 1364
Qy 62 GluGlyThrProArgGlnSerArg 69
Db 1365 CTTGAGACGGAGAGAGCGCGCGG 1388

Search completed: August 14, 2004, 19:16:59
Job time : 152.298 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 832.674 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_450_520

Perfect score: 372

Sequence: 1 TFWRSQVFIDHPVHLKFNVS.....RLLTQEARSLGTRQSRGT 71

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh

-Q=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112216_29287/app query.fasta_1.3519

-DB=EST -Qfmt=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10029020@cgn_1_13135@runat_06082004_112216_29287 -NCFU=6 -ICPV=3

-NO_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_nam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 302 | 81.2 | 870 | 10 | BG036207 |
| 2 | 248.5 | 66.8 | 2627 | 11 | AK050784 |
| 3 | 228 | 61.3 | 701 | 28 | AZ661742 |
| 4 | 220.5 | 59.3 | 3447 | 11 | AK047423 |
| 5 | 202.5 | 54.4 | 594 | 13 | BQ257598 |
| 6 | 201 | 54.0 | 519 | 13 | BQ257598 |
| 7 | 196.5 | 52.8 | 306 | 14 | CF538752 |
| 8 | 194.5 | 52.3 | 2716 | 11 | AK031198 |
| 9 | 186 | 50.0 | 1013 | 12 | BM017187 |
| 10 | 171 | 46.0 | 247 | 9 | AL710528 |
| 11 | 143 | 38.4 | 768 | 29 | BX172779 |
| 12 | 143 | 38.4 | 769 | 29 | BX245366 |
| 13 | 142 | 38.2 | 332 | 14 | N67072 |
| 14 | 139 | 37.4 | 749 | 29 | BX178536 |
| 15 | 117 | 31.5 | 340 | 28 | CC122834 |
| 16 | 83 | 22.3 | 1052 | 9 | AA051165 |
| 17 | 78.5 | 21.1 | 456 | 9 | AI762835 |
| 18 | 75 | 20.2 | 562 | 29 | DR20114T |
| 19 | 74.5 | 20.0 | 547 | 10 | BF498290 |
| 20 | 73 | 19.6 | 600 | 28 | BZ894471 |
| 21 | 73 | 19.6 | 1148 | 12 | BM801011 |
| 22 | 72 | 19.4 | 644 | 28 | AZ083567 |
| 23 | 71 | 19.1 | 426 | 14 | CB274460 |
| 24 | 71 | 19.1 | 433 | 29 | BX179117 |
| 25 | 71 | 19.1 | 593 | 12 | BM355690 |
| 26 | 70.5 | 19.0 | 498 | 10 | BF501431 |
| 27 | 70.5 | 19.0 | 1073 | 28 | BZ559588 |
| 28 | 70 | 18.8 | 348 | 13 | BY046474 |
| 29 | 70 | 18.8 | 446 | 14 | CD453939 |
| 30 | 70 | 18.8 | 511 | 14 | CB447588 |
| 31 | 70 | 18.8 | 592 | 13 | BQ777039 |
| 32 | 70 | 18.8 | 720 | 28 | AQ005704 |
| 33 | 69.5 | 18.7 | 1372 | 13 | BU901558 |
| 34 | 69 | 18.5 | 397 | 14 | N77794 |
| 35 | 69 | 18.5 | 414 | 9 | AI173157 |
| 36 | 69 | 18.5 | 438 | 12 | BM841907 |
| 37 | 69 | 18.5 | 481 | 13 | CA128790 |
| 38 | 69 | 18.5 | 579 | 14 | CD894862 |
| 39 | 68.5 | 18.4 | 377 | 10 | BF423355 |
| 40 | 68.5 | 18.4 | 542 | 12 | BJ203521 |
| 41 | 68 | 18.3 | 369 | 13 | BY294778 |
| 42 | 68 | 18.3 | 476 | 13 | BY031967 |
| 43 | 68 | 18.3 | 503 | 9 | AJ434765 |
| 44 | 68 | 18.3 | 604 | 12 | BU606630 |
| 45 | 68 | 18.3 | 666 | 12 | BI959265 |

ALIGNMENTS

| RESULT | LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE |
|--------|----------|---|-----------|------------|-------------|--------|--|-----------|
| 1 | BG036207 | 602326960F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4428351 5', mRNA sequence. | BG036207 | BG036207.1 | GI:12431132 | EST. | Homo sapiens (human) | |
| | | | | | | | Homo sapiens | |
| | | | | | | | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | |
| | | | | | | | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| | | | | | | | 1 (bases 1 to 870) | |

| | | | | | |
|---|---|--|---|---|--|
| AUTHORS TITLE JOURNAL COMMENT | NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgaps-remail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1AM10178 row: h column: 16 High quality sequence stop: 713. Location/Qualifiers 1. .870 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4428351" /tissue_type="adenocarcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 91" /note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library." | | MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS | 99279253 10349636 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 20499374 11042159 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 2627) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ , Fax: 81-45-503-9216) | |
| | COMMENT cdna library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers 1. .2627 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:D030019H23" /db_xref="MGI:2418637" /db_xref="taxon:10090" /clone="D030019H23" /tissue_type="whole body" | | | | |
| | FEATURES source | | | | |
| | US-10-029-020-14_COPY_450_520 (1-71) x BG036207 (1-870) | | | | |
| ORIGIN | Alignment Scores: Pred. No.: 2.79e-30 Length: 870 Score: 302.00 Matches: 69 Percent Similarity: 92.00% Conservative: 0 Best Local Similarity: 92.00% Mismatches: 1 Query Match: 81.18% Indels: 5 DB: 10 Gaps: 0 | | TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS | | |
| | QY 1 ThrPheTtpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20 | | | | |
| | Db 478 ACTTCTTGAGATCTCAAGTGTTCATAGACCATCTGTGCAATTCGAATTCATGTGCT 537 | | | | |
| | QY 21 LeuGlyLysAlaAlaLeu--ValGlyIleTyrGlyArgLysGlyLeuProSerHis- 39 | | | | |
| US-10-029-020-14_COPY_450_520 (1-71) x BG036207 (1-870) | 538 CTGGGAAGGCAGCCCTGGGGTTGGCATTTATGGCAGAAAGCCCTCCCTTCACATT 597 | | TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS | | |
| | QY 40 ThrGlnPheAspPheValGluLeuLeuAspGlyArgArgLeu-LeuThrGlnGluAlaAr 59 | | | | |
| | Db 598 ACACAGTTTGACTTTGTGAGCTGCTGGATGGCCGAGGCTCCCTAACCCAGAGCGCG 657 | | | | |
| | QY 59 gSerLeuGlu-GlyThrProArgGlnSerArgGly 70 | | | | |
| RESULT 2 AK050784 LOCUS DEFINITION | 658 GAGCCTAGAGGGGAGCCCGCCGACCTCTTCGGGG 692 | | COMMENT Mus musculus 9 days embryo whole body cDNA, RIKEN full-length enriched library, clone:D030019H23 product:odd Oz/ten-m homolog 3 (Drosophila), full insert sequence. AK050784.1 GI:26094117 AK050784.1 HTC 20-SEP-2003 Mus musculus (house mouse) HTC; CAP trapper. KEYWORDS Mus musculus (house mouse) Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) | | |
| | ACCESSION VERSION KEYWORDS SOURCE ORGANISM | | | | |
| | REFERENCE AUTHORS TITLE JOURNAL | | | | |
| | | | | | |


```

/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="9 days embryo"
1. .2627
misc_feature
/notes="odd Oz/ten-m homolog 3 (Drosophila)
(MGB|GI:1345183, GB|NM_011857, evidence: BLASTN, 99%,
match=1753)"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      2.7e-22      Length:      2627
Score:          248.50      Matches:      48
Percent Similarity: 80.28%      Conservative: 9
Best Local Similarity: 67.61%      Mismatches: 11
Query Match:    66.80%      Indels:      3
DB:             11          Gaps:        1

```

US-10-029-020-14_COPY_450_520 (1-71) x AK050784 (1-2627)

```

Qy 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 777 TTCGGAGATCGCAGCTCTTTATTGATCAGCCACAGTTCTTAAGTTCAACATCTCTT 836
Qy 22 GlyLysAlaAlaLeuValGlyIleTyTGlyArgLysGlyLeuProSerHisThrGln 41
Db 837 CAGAAGGATGCAATTCATCGGAGTGTACGGCGGAAGGCTTACCGCTTCCCATCTCAG 896
Qy 42 PheAspPheValGluLeuLeuAspGlyArgGlnLeuThrGlnGluAlaArgSerLeu 61
Db 897 TAGCACTTGTGGNACTACTGATGGTAGCAGGTTAATTCGAGAGAGACGCGGAACCTG 956
Qy 62 -----GluGlyThrProArgGlnSerArg 69
Db 957 GTGAGTCGCGAAGAGAGCGCGGCGGCGGCGAGAG 989

```

RESULT 3

```

AZ661742
LOCUS 1M0540B02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0540B02 R, genomic survey sequence.
ACCESSION AZ661742
VERSION AZ661742.1 GI:11798888
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 701)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhauser,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0540 row: B column: 02
Seq primer: CACACAGGAACGATGATGACC
Class: plasmid ends
High quality sequence stop: 701.
Location/Qualifiers
1. .701
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

```

FEATURES source

```

/clone="UUGC1M0540B02"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 [gi|4732114|gb|AF129072.1], a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      2.39e-20      Length:      701
Score:          228.00      Matches:      45
Percent Similarity: 83.93%      Conservative: 2
Best Local Similarity: 80.36%      Mismatches: 9
Query Match:    61.29%      Indels:      0
DB:             28          Gaps:        0

```

US-10-029-020-14_COPY_450_520 (1-71) x AZ661742 (1-701)

```

Qy 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 128 ACGTTCTGGAGATCTCAGGTTTCATAGACCCCTGTACACCTGAAGTTCAATGTGCT 187
Qy 21 LeuGlyLysAlaAlaLeuValGlyIleTyTGlyArgLysGlyLeuProSerHisThr 40
Db 188 CTGGGGAAGCAGCTCTGGTTGGCATTATGGCAGAAAGGCTTCTCTCTCCCATCT 247
Qy 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlnLeuThrGln 56
Db 248 CAGGTAACCTAGACTCCACCCCTGAGGGACCTCTTGATCTTAACCCAG 295

```

RESULT 4

AK047423

LOCUS

DEFINITION

AK047423 Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:B930059M16 product:odd Oz/ten-m homolog 1 (Drosophila), full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

AK047423 3447 bp mRNA linear HTC 20-SEP-2003
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:B930059M16 product:odd Oz/ten-m homolog 1 (Drosophila), full insert sequence.

AK047423.1 GI:26338765

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

```
11042159
3
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishihata, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3447)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES
Location/Qualifiers
1..3447
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:B930059M16"
/db_xref="MGI:2412967"
/db_xref="taxon:10090"
/clone="B930059M16"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
862..2472
/notes="unnamed protein product; odd Oz/ten-m homolog 1
(Drosophila) (MGD|MGI:1345185, GB|NM_011855, evidence:
BLASTN, 99%, match=3037)
putative"
/codon_start=1
/protein_id="BAC33054.1"
CDS
/db_xref="GI:26338766"
/translation="MEQTCKPYQPLSKVKHEMDLAYTSSSEDESGRKPQSPNSRE
TLHEYDCLRNVSQSRKDKVEKSTQELFCETPTLCGVTDMHVSRRHQYLE
MGSDVTEGASPPHALFMWIRGKSHSSCLSSRANSALSLTDTHERKSDGNG
FKSPVCCDMEAFADSAQDMQSPHNOFTPRPLPPPPHACTCAKPPPTVDSLOR
RSMYTSQSPAPAPPTSDQSVLHNSWVLSNIPLETRHFLFKHGSSSAIFSA
SONYPTNTVYSPPPRPLRSTFSPAPFTKPYRCNWKCTALSATITVUALLL
AVIVLHLGTWLOLPGVQIYANGISNGNPTESMDTYSPIGGRVSKDKKVKFQ
RAIDTGSDVIGAOVMOTIPGLFWFOITTHPIYLFKNISLAKDSLLGIYGRNI
THQTFDFVLMKQKLVKQSKSDDIQHSRNLILTSLOETGFIYMDQGPWYLA
Y
NDGKKMEQVFLITAIKGLPCLLEPFTV"
ORIGIN
Alignment Scores:
Pred. No.: 2,6e-18 Length: 3447
Score: 220.50 Matches: 37
Percent Similarity: 83.82% Conservativeness: 20
Best Local Similarity: 54.41% Mismatches: 8
Query Match: 59.27% Indels: 3
DB: 11 Gaps: 1
US-10-029-020-14_COPY_450_520 (1-71) x AK047423 (1-3447)
QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 2104 TTCTGGCGTTTCCAGATTACTATCCACCATCTATATCTGAAATTCATTTCTTTA 2163
QY 22 GlyLysAlaLeuValGlyLeuValGlyLeuValGlyLeuValGlyLeuValGlyLeu 41
Db 2164 GCCAAGGACTCTCTTCTGGGAATTATGCGAGAGAGAAACATCCACCTACACACTCAG 2223
QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer--- 60
Db 2224 TTTGATTTCTGGAATTAATGATGCGCAACAACTGGTAAACAGGACTCTAAGAGCTCA 2283
QY 61 -----LeuGluGlyThrProArg 66
Db 2284 GATGACATTCAGCACTCCCAAGG 2307
RESULT 5
BQ257598 594 bp mRNA linear EST 06-MAY-2002
NISC Kp04a02.q2 Baker mouse embryo e7.5 Mus musculus cDNA clone
IMAGE:5408163, mRNA sequence.
ACCESSION BQ257598
VERSION BQ257598.1 GI:20458351
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 594)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: J. Baker (Stanford University)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
infoimage.llnl.gov
MGI:1844899
Plate: LLAM12041 row: B column: 4
Seq primer: Sp6 primer.
Location/Qualifiers
1..594
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"
```


/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BNAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 1.37e-16 Length: 306
 Score: 196.50 Matches: 39
 Percent Similarity: 81.03% Conservative: 8
 Best Local Similarity: 67.24% Mismatches: 8
 Query Match: 52.92% Indels: 3
 DB: 14 Gaps: 1

US-10-029-020-14_COPY_450_520 (1-71) x CF538752 (1-306)

Qy 15 LeuYsPheAenValSerLeuGlyLysAlaAlaLeuValGlyLeuTyrGlyArgLysGly 34
 Db 9 CTTAAGTTCAACATCTCTCTTCAGAAAGGATGCAATTCATCGGAGTGTACGGCGGAAGGGC 68

Qy 35 LeuProProSerHisThrGlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeu 54
 Db 69 TTACCGCTTCCCATCTACGATCAGTACGACTTTGTGGAACTACTCGATGGTAGCAGGTTAAIT 128

Qy 55 ThrGlnGluAlaArgSerLeu-----GluGlyThrProArgGlnSerArg 69

Db 129 CGGAGAGCAGCGGACCTGTGTGAGTCGAGCGGCGGCGGCGGAGAGA 182

RESULT 8
 AK031198 LOCUS
 DEFINITION Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length enriched library, clone:5930427L24 product:odd Oz/ten-m homolog 2 (Drosophila), full insert sequence.

ACCESSION AK031198.1 GI:26082195
 VERSION AK031198.1
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL 99279253
 MEDLINE 10349636
 PUBMED

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL 20499374
 MEDLINE 11042159
 PUBMED

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Toga, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

TITLE

JOURNAL
 MEDLINE
 PUBMED

REFERENCE

AUTHORS

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

TITLE

JOURNAL

REFERENCE

AUTHORS

The RIKEN Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 409, 685-690 (2001)

TITLE

JOURNAL

REFERENCE

AUTHORS

The RIKEN Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

TITLE

JOURNAL

REFERENCE

AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN
 Genomic Sciences Center and Genome Science Laboratory in RIKEN
 Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

TITLE

JOURNAL

REFERENCE

AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN
 Genomic Sciences Center and Genome Science Laboratory in RIKEN
 Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN
 Genomic Sciences Center and Genome Science Laboratory in RIKEN
 Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source

1..2716
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:5930427L24"
 /db_xref="MGI:2394397"
 /db_xref="taxon:10090"
 /clone="5930427L24"
 /tissue_type="forelimb"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="13 days embryo"
 1..2716
 /note="odd Oz/ten-m homolog 2 (Drosophila)
 (MGDI:1345184, GB|NM_011856, evidence: BLASTN, 99%, match=1708)"

misc_feature

1..2716
 /note="odd Oz/ten-m homolog 2 (Drosophila)
 (MGDI:1345184, GB|NM_011856, evidence: BLASTN, 99%, match=1708)"

ORIGIN

Alignment Scores:
 Pred. No.: 6.34e-15 Length: 2716
 Score: 194.50 Matches: 38
 Percent Similarity: 70.59% Conservative: 10
 Best Local Similarity: 55.88% Mismatches: 15
 Query Match: 52.28% Indels: 5
 DB: 11 Gaps: 1

US-10-029-020-14_COPY_450_520 (1-71) x AK031198 (1-2716)


```

Db      216  TATGACTTCATGGAACCTCTGGACGGGAG 245
          ::::::::::::::::::::
BX172779      768 bp      DNA      linear      GSS 28-JAN-2003
Danio rerio genomic clone DKEY-176L10, genomic survey sequence.
ACCESSION      BX172779
VERSION      BX172779.1  GI:28004484
KEYWORDS      GSS.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE      1 (bases 1 to 768)
AUTHORS      Humphray,S.J., Huckle,E. and Durham,J.I.
TITLE      Direct Submission
JOURNAL      Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphray@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 176L10. 176L10
is part of the Daniokey BAC library created by R. Piasterk and N.V.
keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source           1..768
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-176L10"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
ORIGIN
Alignment Scores:
Pred. No.:           1.01e-08      Length:      768
Score:              143.00         Matches:     29
Percent Similarity: 69.39%         Conservative: 5
Best Local Similarity: 59.18%      Mismatches:  11
Query Match:       38.44%          Indels:      4
DB:                29              Gaps:         1
US-10-029-020-14_COPY_450_520 (1-71) x BX172779 (1-768)
Qy      2  PheTrpArgSerGlnValPheileAspHisProValHisLeuLysPheAsnValSerLeu 21
          ::::::::::::::::::::
Db      702  TTCGGAGATCCCTCATGCACCTTTACACGCTCAGTCTCCATCAAGTTCAACATCTCCTTG 643
          ::::::::::::::::::::
Qy      22  GlyLysAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHis----- 39
          ::::::::::::::::::::
Db      642  GGCAAGGATGCACGTGTTTGGGGTTTACATTCGAAAGGGCTGCCCGCTCCATGCACAG 583
          ::::::::::::::::::::
Qy      40  -----ThrGlnPheAspPheValGlu 46
          ::::::::::::::::::::
Db      582  GTGCGCACACAGCTTTCTATTACAGAA 556
          ::::::::::::::::::::
RESULT 12
BX245366/c
LOCUS      BX245366
DEFINITION  Danio rerio genomic clone DKEY-230G8, genomic survey sequence.
ACCESSION      BX245366
VERSION      BX245366.1  GI:28167700
KEYWORDS      GSS.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE      1 (bases 1 to 769)
AUTHORS      Humphray,S.J., Huckle,E. and Durham,J.I.
TITLE      Direct Submission
JOURNAL      Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome

```

```

source
1. .392
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3893111"
/db_xref="taxon:9606"
/clone="IMAGE:285994"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Morton Fetal Cochlea"
/notes="Organ: ear; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned unidirectionally. Primer: Oligo dt. Fetal cochlea, normal. 3x of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP XR vector. Library constructed by N. Robertson, C. Morton. ~5' adaptor sequence: 5' GAATTCGGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTT 3'"

ORIGIN
Alignment Scores:
Pred. No.: 5.15e-09 Length: 392
Score: 142.00 Matches: 29
Percent Similarity: 85.00% Conservative: 5
Best Local Similarity: 72.50% Mismatches: 6
Query Match: 38.17% Indels: 1
DB: 14 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x M67072 (1-392)

QY 2 PheTtpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 177 TTTCTGAGATCAGCTTTCATTGATCAGCCACAGTTTCTTAATTCATATCTCTCT 118

QY 22 GlyLysAlaAlaLeuValGlyIleTyTGlyArgLysGlyLeuProSerHisThrGln 41
Db 117 CAGAGGATGCATGATGAGTATATGGCGGAGAG-TTACCGCTTCCCATACTCAG 59

RESULT 14
BX178536/c
LOCUS
DEFINITION
Danio rerio genomic clone DKEY-172P5, genomic survey sequence.
ACCESSION
BX178536
VERSION
BX178536.1 GI:28010314
KEYWORDS
GSS.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 749)
Humphray, S.J., Huckle, E. and Durham, J.L.
Direct Submission
Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 172P5. 172P5 is
part of the DanioKey BAC Library created by R. Plasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/projects/D_rerio/

FEATURES
source
1. .749
Location/Qualifiers
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-172P5"
/tissue_type="Testis"
/notes="vector pindigoBAC-536"

ORIGIN
Alignment Scores:
Pred. No.: 3.41e-08 Length: 749
Score: 139.00 Matches: 26
Percent Similarity: 75.00% Conservative: 4

```

```

Best Local Similarity: 65.00% Mismatches: 10
Query Match: 37.37% Indels: 0
DB: 29 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x BX178536 (1-749)

QY 2 PheTtpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 382 TTTCTGAGGTCCTCTTACCTACCTCGCCAGCCACACTTCTCCTCAAGTTCAACATCTCCTG 323

QY 22 GlyLysAlaAlaLeuValGlyIleTyTGlyArgLysGlyLeuProSerHisThrGln 41
Db 322 GGCAGGACGGCTGTGGGGTCTACATCGCGCAAGGCTGCTCTTCCACGCTCAG 263

RESULT 15
CC122834/c
LOCUS
DEFINITION
NDL.45P17, Notre Dame Liverpool Aedes aegypti genomic clone
ACCESSION
CC122834
VERSION
CC122834.1 GI:29991889
KEYWORDS
GSS.
SOURCE
Aedes aegypti (yellow fever mosquito)
ORGANISM
Aedes aegypti
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;
Stegomyia.
REFERENCE
1 (bases 1 to 340)
Loftus, B., Shetty, J., Knudson, D. and Severson, D.
TITLE
BAC end sequencing of Aedes aegypti
JOURNAL
Unpublished (2003)
COMMENT
Other_GSSs: NDL.45P17.SP6
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org
Library was provided by David Severson
Seq primer: T7
Class: BAC ends.

FEATURES
source
1. .340
Location/Qualifiers
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="liverpool"
/db_xref="taxon:7159"
/clone="NDL.45P17"
/clone_lib="Notre Dame Liverpool"
/notes="Vector: pECBAC1; Site 1: Hind III; The library was
prepared from whole body tissue of newly hatched L1 larvae
by David Severson at the University of Notre Dame and
Hongbin Zhang"

ORIGIN
Alignment Scores:
Pred. No.: 1.05e-05 Length: 340
Score: 117.00 Matches: 21
Percent Similarity: 61.54% Conservative: 11
Best Local Similarity: 40.38% Mismatches: 20
Query Match: 31.45% Indels: 0
DB: 28 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x CC122834 (1-340)

QY 2 PheTtpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 250 TTTTGGACGCGAGATTTAGTAACAAACATCTCGCATTCATTAATTTCAATTTTA 191

QY 22 GlyLysAlaAlaLeuValGlyIleTyTGlyArgLysGlyLeuProSerHisThrGln 41
Db 190 CCGTGGGGTGCCAACTTTGCTGTGTACGTCGACGAATGTAGGACCTAGCATACGCAA 131

```

Qy 42 PheaspPheValGluLeuLeuAspGlyArgArgLeu 53
Db 130 TATGATTTGTAGAATTATATAAAGGTGGAAGGATA 95

Search completed: August 14, 2004, 18:02:47
Job time : 842.674 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 4691.12 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_1_400

Perfect score: 2143

Sequence: 1 MDVKRPYRSLRRDRDAER.....ETEDTASSWFPVPTDVSLLP 400

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10029020/runat_06082004_112216_29287/app.query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029020 @cgn 1.13135 @runat 06082004_112216_29287 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:.*
1: em_estba:.*
2: em_esthum:.*
3: em_estinh:.*
4: em_estmu:.*
5: em_estov:.*
6: em_estpl:.*
7: em_estro:.*
8: em_htc:.*
9: gb_estl:.*
10: gb_est2:.*
11: gb_htc:.*
12: gb_est3:.*
13: gb_est4:.*
14: gb_est5:.*
15: em_estfun:.*
16: em_estom:.*
17: em_gss_hum:.*
18: em_gss_inv:.*
19: em_gss_pln:.*
20: em_gss_vrt:.*
21: em_gss_fun:.*
22: em_gss_mam:.*
23: em_gss_mus:.*
24: em_gss_pro:.*
25: em_gss_rod:.*
26: em_gss_phg:.*
27: em_gss_vrl:.*
28: gb_gss1:.*

29: gb_gss2:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|----------|---------------------|
| C 1 | 936 | 43.7 | 620 | 14 | CB532245 | CB532245 756697 MA |
| 2 | 936 | 43.7 | 620 | 14 | CB532612 | CB532612 757081 MA |
| 3 | 884 | 41.3 | 566 | 13 | BU072782 | BU072782 im47902.Y |
| 4 | 781 | 36.4 | 625 | 13 | BU200362 | BU200362 604158648 |
| 5 | 737 | 34.4 | 878 | 13 | BU474934 | BU474934 603364850 |
| 6 | 722.5 | 33.7 | 3447 | 11 | AK047423 | AK047423 Mus muscu |
| 7 | 699 | 32.6 | 884 | 13 | BQ735622 | BQ735622 AGENCOURT |
| 8 | 692 | 32.3 | 789 | 13 | BU474512 | BU474512 603760992 |
| 9 | 687.5 | 32.1 | 887 | 13 | BU119163 | BU119163 603143046 |
| 10 | 685 | 32.0 | 402 | 14 | CB544750 | CB544750 ANGNNUC:T |
| 11 | 659 | 30.8 | 4556 | 11 | AK034286 | AK034286 Mus muscu |
| 12 | 596.5 | 27.8 | 490 | 9 | AL922332 | AL922332 AL922332 |
| 13 | 568 | 26.5 | 870 | 10 | BG036207 | BG036207 602326960 |
| 14 | 531.5 | 24.8 | 511 | 29 | AY405291 | AY405291 Homo sapi |
| 15 | 531.5 | 24.8 | 511 | 29 | AY405292 | AY405292 Pan trogl |
| 16 | 529.5 | 24.7 | 511 | 29 | AY405293 | AY405293 Mus muscu |
| 17 | 524.5 | 24.5 | 534 | 10 | BF944080 | BF944080 QV3-NN020 |
| 18 | 509 | 23.8 | 2716 | 11 | AK031198 | AK031198 Mus muscu |
| 19 | 502.5 | 23.4 | 880 | 13 | BF760712 | BF760712 BX760712 |
| 20 | 482.5 | 22.5 | 568 | 10 | BF953105 | BF953105 QV3-NN020 |
| 21 | 470 | 21.9 | 2627 | 11 | AK050784 | AK050784 Mus muscu |
| 22 | 456.5 | 21.3 | 643 | 10 | BG657984 | BG657984 BB657984 |
| 23 | 448 | 20.9 | 717 | 10 | BG654584 | BG654584 BB654584 |
| 24 | 438.5 | 20.5 | 593 | 29 | CE756467 | CE756467 tigr-gss- |
| 25 | 433 | 20.2 | 284 | 10 | BF388169 | BF388169 UI-R-CAI- |
| 26 | 417 | 19.5 | 259 | 28 | CC178145 | CC178145 XC549 Bay |
| 27 | 409 | 19.1 | 825 | 13 | BU253342 | BU253342 603747714 |
| 28 | 388.5 | 18.1 | 843 | 14 | CF225019 | CF225019 AGENCOURT |
| 29 | 381 | 17.8 | 864 | 13 | BU172491 | BU172491 AGENCOURT |
| 30 | 374 | 17.5 | 397 | 13 | BY010856 | BY010856 BY010856 |
| 31 | 374 | 17.5 | 704 | 13 | BY733523 | BY733523 BY733523 |
| 32 | 367 | 17.1 | 643 | 13 | BY723994 | BY723994 BY723994 |
| 33 | 343.5 | 16.0 | 608 | 9 | AL955817 | AL955817 AL955817 |
| 34 | 330.5 | 15.4 | 707 | 13 | BU229930 | BU229930 603947544 |
| 35 | 329.5 | 15.4 | 801 | 13 | BU451304 | BU451304 603216940 |
| 36 | 295 | 13.8 | 461 | 13 | BU471305 | BU471305 603761665 |
| 37 | 282.5 | 13.2 | 404 | 29 | CG514853 | CG514853 OST68836 |
| 38 | 282.5 | 13.2 | 751 | 13 | BU461850 | BU461850 603773816 |
| 39 | 273.5 | 12.8 | 1060 | 29 | CNS04V97 | AL308788 Tetraodon |
| 40 | 272.5 | 12.7 | 745 | 29 | CE434866 | CE434866 tigr-gss- |
| 41 | 257.5 | 12.0 | 386 | 9 | AA702693 | AA702693 x190904.s |
| 42 | 256.5 | 12.0 | 674 | 28 | B2111628 | B2111628 CH230-230 |
| 43 | 248 | 11.6 | 474 | 9 | AL753786 | AL753786 crl6e09.x |
| 44 | 247.5 | 11.5 | 600 | 13 | BU921705 | BU921705 60931-12 M |
| 45 | 247.5 | 11.5 | 644 | 10 | BB657539 | BB657539 BB657539 |

ALIGNMENTS

RESULT 1
CB532245/c
LOCUS CB532245 620 bp mRNA linear EST 16-MAY-2003
DEFINITION 756697 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
ACCESSION CB532245
VERSION CB532245.1 GI:29395750
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 620)

Db 92 AACACGAGACCGATCACCGCGCGCTGCAGAACATTCTCGACTCCGACACCGCG 33

QY 181 ProProLeuSerHisAlaHisThrPro 189
|||||

Db 32 CTTCCGCTCTCGCATGCCACACCCCC 6
|||||

RESULT 2
CB532612 620 bp mRNA linear EST 16-MAY-2003
LOCUS 757081 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION CB532612
ACCESSION CB532612.1 GI:29396493
VERSION
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 620)
AUTHORS Smith,T.P.L., Roberts,A.J., Echterkamp,S.E., Chitko-McKown,C.G.,
Wray,J.E. and Keeler,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: LAM8014 row: K column: 19
Seq primer: GTAATACGACTCACTATAGG.
FEATURES
source
1..620
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

ORIGIN
Alignment Scores:
Pred. No.: 2.05e-58 Length: 620
Score: 936.00 Matches: 174
Percent Similarity: 95.77% Conservatve: 7
Best Local Similarity: 92.06% Mismatches: 8
Query Match: 43.68% Indels: 0
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x CB532612 (1-620)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
|||||

Db 49 ATGGAGCTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 108
|||||

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
|||||

Db 109 CGCTACACCGAGCTCGTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 168
|||||

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
|||||

Db 169 TCCAGGAGACCTCAAGGCTTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 228
|||||

QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
|||||

Db 229 AAGGACTGGTACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 288
|||||

AUTHORS Smith,T.P.L., Roberts,A.J., Echterkamp,S.E., Chitko-McKown,C.G.,
Wray,J.E. and Keeler,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: LAM8014 row: K column: 19
Seq primer: TAGAAGGACGACGTCGAGG.
FEATURES
source
1..620
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

ORIGIN
Alignment Scores:
Pred. No.: 2.05e-58 Length: 620
Score: 936.00 Matches: 174
Percent Similarity: 95.77% Conservatve: 7
Best Local Similarity: 92.06% Mismatches: 8
Query Match: 43.68% Indels: 0
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x CB532245 (1-620)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
|||||

Db 572 ATGGAGCTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 513
|||||

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
|||||

Db 512 CGCTACACCGAGCTCGTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 453
|||||

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
|||||

Db 452 TCCAGGAGACCTCAAGGCTTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 393
|||||

QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
|||||

Db 392 AAGGACTGGTACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 333
|||||

QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuThrArgThrAspIle 100
|||||

Db 332 CGCGAGCTGGGGTGGGGAGGAGTGACACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 273
|||||

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
|||||

Db 272 GGCTCTCCCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 213
|||||

QY 121 ThrValLeuSerProGluHisProValArgLeuTyrPglyArgSerThrArgSerGlyArg 140
|||||

Db 212 GCGGCATCTGTCCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 153
|||||

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
|||||

Db 152 AGCTCTGCTGTCCAGCGGGGCACTCCCACTCACTCACTCACTCACTCACTCACTCACTCA 93
|||||

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
|||||

QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
 Db 289 CGCGAGCTGGGGTGGGGAGGTGACACCCCGGACGGGAGCGGTGTACCGACGGACATC 348
 QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
 Db 349 GGCCTGGCCCACTCGGGCTACTCCCTGGGGCCGAGCTCTGAGGCGGAGCTGGAGCGGAC 408
 QY 121 ThrValLeuSerProGluHisProValArgLeuThrPpGlyArgSerThrArgSerGlyArg 140
 Db 409 GCGGCATGTCCCGGAGACCCCGTGGCGGTGTGGGGCCGACGACGCGGTCTGAGGACG 468
 QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
 Db 469 AGCTCTCGCTGTCCAGCGGGCCCACTCCCACTCTCAGCTCACTGACACGAGACGAG 528
 QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
 Db 529 AACAGGAGACCGGATCAACCGGGCGGCTGCAGAACCACTTCGACTCCGGACACCGCG 588
 QY 181 ProProLeuSerHisAlaHisThrPro 189
 Db 589 CCTCGGCTCTCGCATGCCCAACCCCC 615

RESULT 3
 BU072782
 LOCUS
 DEFINITION
 im47g02.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6038211 5'
 similar to TR:014667 O14667 PRO-NEUREGULIN-1, GAMMA ISOFORM
 PRECURSOR ; mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
 Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
 Hillier,L., Maria,M., Pape,D., Wyllie,T., Martin,J., Blistain,A.,
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
 Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R.,
 Williams,T., Jackson,Y. and Bowers,Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)

TITLE
 JOURNAL
 COMMENT
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohph.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@im.wustl.edu)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 339.
 Location/Qualifiers
 1..566
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6038211"
 /tissue type="Purified pancreatic islet"
 /lab host="DH10B"
 /clone lib="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
 Noti; Site_2: XhoI; cDNA made by oligo-dT priming.
 Size-selected on agarose gel. Average insert size ~1kb. 5'

FEATURES
 source

XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,12e-54 Length: 566
 Score: 884.00 Matches: 162
 Percent Similarity: 95.86% Conservative: 0
 Best Local Similarity: 95.86% Mismatches: 7
 Query Match: 41.25% Indels: 0
 DB: 13 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x BU072782 (1-566)

QY 165 AspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProProLeuSer 184
 Db 57 GATCATCCGGCGGCTGTGAGAACACCGCGGCTCCGGACGCGCGCGCGCTCTCG 116
 QY 185 HisAlaHisThrProAsnGlnHisHisAlaSerIleAsnSerLeuAsnArgGlyAsn 204
 Db 117 CAGCCCAACACCCCAACACCGACCGCGCTCCATTAACTCCCTGAACCGGGGCAAC 176
 QY 205 PheThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluPro 224
 Db 177 TTCACGCGGAGGAGCAACCCACCGCGGCCCCACGGACCACTCGCTCCGGAGAGCCC 236
 QY 225 ProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsn 244
 Db 237 CCTGGCGGGCGGCGCCAGGAGCGCTGCCACCGCCAGGAGAACTGGCTCAACAGCAAC 296
 QY 245 IleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsn 264
 Db 297 ATCCCCCTGGAGACCAAGAAACCTAGGCAAGACGCCATTCTAGGAGCAATTGCAAGCAAC 356
 QY 265 LeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHis 284
 Db 357 CTCATTGAGATGACATTTCTGGCGCTCCCGCCATGATGGGCTTACAGTACGGGAC 416
 QY 285 PheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrPro 304
 Db 417 TTCTCTTCAAGCTGGAGGCACCTCCCGCTCTTCTGCACCAACATCACAGGGTACCCA 476
 QY 305 LeuThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAla 324
 Db 477 CTGAGTCCAGCACAGTGACTCTCCCTCCGCGCCGACCCCTGGCGGACGACCACTTTCG 536
 QY 325 ArgProAlaPheAsnLeuLysLysPro 333
 Db 537 CGGGCGGNGCTTTAACCCTCAAGAAGCCT 563

RESULT 4
 BU200362
 LOCUS
 DEFINITION
 604158648F1 CSEQCHN03 Gallus gallus cDNA clone CHEST1010p11 5',
 mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Gallus gallus (chicken)
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (Bases 1 to 625)
 Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.U.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 2235534

12445392
 Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

1. 625
 Location/Qualifiers
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CHEST1010p11"
 /tissue_type="whole embryo"
 /dev_stage="20-21"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN03"
 /note="Organ: whole embryo; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

FEATURES
 source
 Alignment Scores:
 Pred. No.: 4.25e-47 Length: 625
 Score: 781.00 Matches: 151
 Percent Similarity: 91.62% Conservative: 13
 Best Local Similarity: 84.36% Mismatches: 13
 Query Match: 36.44% Indels: 2
 DB: 13 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x BU200362 (1-625)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
 Db 93 ATGGATGTAAGAAGAAACCGATATCGATCTCTGACTCGCGCGGACACGAGCGC 152

QY 21 ArgTyrThrSerSerSerAlaSerSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
 Db 153 CGCTACACCAAGCTCTTCAGCGAGAGTGAAGAGCAGCAGCAGCGCTCTCAGAAAGTCCATATAGC 212

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
 Db 213 TCCAGTGAACCTTGAAGGCTTATGATCAAGACTCCAGCTTGCCTGACCTACGCAATCGGTC 272

QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaSerPheThrLeu 80
 Db 273 AAGAATATGTCACACGAGGCTGATGATTCGCGAGCAGGAGCAACTTCTCTTTG 332

QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
 Db 333 CGGAGCTGGTCTTGAAGATGTGACTCCACCCATGGGACTTTGTACCGGATGATATT 392

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
 Db 393 GGGCTGCTCCTACTGTGGCTACTCCATCAGCGCTGCTCAGATCCGACACGAGGCGGAT 452

QY 121 ThrValLeuSerProGluHis-ProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
 Db 453 GTGGTCAATGTCCTAGCATCCCTGTGAGGCTCTGGGAGCGCAACCAATCCGAGC 512

140 gSerSerCysLeuSerSerArgAlaSerSerAsnLeuThrLeuThrAspThrGluHisG1 160
 Db 513 CACCTCTCTCTTCTCGAGCGGCGCAACTCCAACTCAGCTCAGCTCAGCGAGCAGCA 572

QY 160 uAsn-ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArg 177
 Db 573 GAATTACCGAAACTGATCATCTCCAAATCTTCAAAATCATTCAAGACTCCGA 625

RESULT 5
 BU474934 878 bp mRNA linear EST 30-NOV-2002
 LOCUS 603364850F1 CSEQREN21 Gallus gallus cDNA clone CHEST262C14 5', mRNA
 sequence.
 ACCESSION BU474934
 VERSION BU474934.1 GI:25968511
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 878)
 Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 JOURNAL 2235534
 MEDLINE 12445392
 PUBMED
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source
 Location/Qualifiers
 1. 878
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Layer"
 /db_xref="taxon:9031"
 /clone="CHEST262C14"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="CSEQREN21"
 /note="Organ: ovary; Vector: pBluescript II KS(+); Site 1:
 EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.13e-43 Length: 878
 Score: 737.00 Matches: 141
 Percent Similarity: 92.73% Conservative: 12
 Best Local Similarity: 85.45% Mismatches: 11
 Query Match: 34.39% Indels: 1
 DB: 13 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x BU474934 (1-878)

Cy 1 MetAspVallyLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
 Db 257 ATGATGTAAAGAAAGAAACCGTATCGATCTCTGACTCGCGCCGCGACACGAGCGGC 316
 Qy 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
 Db 317 CGCTACACCACTCTTCAGCGGAGAGTGGAGCAGCAAGGCTCTCTCAGAGCTCTATAGC 376
 Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
 Db 377 TCCAGTGAGACCCCTGAAGGCTTATGATCAAGACTCCAGGTTGACCTACAGCAATCGGCTC 436
 Qy 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhenThrLeu 80
 Db 437 AAGACATGGTGACACGAGGCTGTAGTAATCTGCCAGCAGGAGCACTTCTTTG 496
 Qy 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
 Db 497 CGGAGCTGGGTCTTGAAGATGTGACTCCCAACCCATGGGACTTGTACCGGACTGATATT 556
 Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
 Db 557 GGGCTGCTCCTCAGTGGCTACTCATCAGCGCTGGCTCAGATGCGGACACGAGGCGGAT 616
 Qy 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
 Db 617 GTGGTCATGTCACCTGAGCATCTCTGTGAGGCTCTGGGAGCGACACCAATCGGACGC 676
 Qy 141 SerSerCysLeuSerSerArgAlaAsnSer-AsnLeuThrLeuThrAspThrGluHisG1 160
 Db 677 AGTCTCTGCTGTGCGAGCGGGCACTCCAAACCTCACCCTCAGCAGCAGGACGCA 736
 Qy 160 uAsnThrGluThr 164
 Db 737 GAATACGGAACCT 749

RESULT 6
 AKO47423
 LOCUS
 DEFINITION
 Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length
 enriched library, clone:B930059M16 product:odd Oz/ten-m homolog 1
 (Drosophila), full insert sequence.
 AKO47423
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 TITLE
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 PUBMED
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sasaki, N., Carninci, P.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 PUBMED
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 3447)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, K., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN,
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.go.jp/
 URL:http://fantom.gsc.riken.go.jp/
 Location/Qualifiers

FEATURES
 Source

1..3447
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6J"
 /db_xref="FANTOM DB:B930059M16"
 /db_xref="MGI:2412967"
 /db_xref="taxon:10090"
 /clone="B930059M16"
 /tissue_type="cerebellum"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="10 days neonate"
 862..862
 /note="unnamed protein product; odd Oz/ten-m homolog 1
 (Drosophila) (MGI:MGI:1345185, GB|NM_011855, evidence:
 BLASTN, 99%, match=3037)
 putative"
 /codon_start=1
 /protein_id="BAC33054.1"
 /db_xref="GI:26338766"

CDS

/translation="MEQTDCKPYOLSKVHEMDLAYTSSDSESDGRKPKQSFNSR
 TLHYNQELRRNYSQSKRDKPKSTIEFCETPTLCGYHTDMHSVRHGYLE
 MGSVDVETEGAGSPDHALRWIRGMKSHSSCLSSALSALSTDDTHERKSGENG
 FKSPVCCMEAPADSAQMSQSPHNPQTFRLPPPPPPHACTCARPKPTVLSIOR
 RSMTRQSPSPAPAPTSTQDSVHLHNSWLNNSIPLERHFLFKHGSGSSAIFSA
 SQNPLTNSNTVYSPPRPLPRSTFRPAFTNKVRCCKWKTALSAITATVTLALL
 AYVIVHLGLTQWLOPVGQIVANGISNGNPGTESMDTYISPIGRVSDSKSEKKVFKG
 RAIDTGEVDIGAQMQLIPGLFWRFQITIIHHPILKFNLSLADGLIGYGRNRIP
 THTQDFVKLMDGKQLVKQDSKSDDIQHSRPNLILTSLETQTFIEYMDQGPWLAFY

Qy 267 GluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheIeu 286
Db 1639 -----CATTCCTG 1647
Qy 287 PheIysProGly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 1648 TTCAACATGATGCTGTTCTTCTGCTATCTTCTGTCAGCCAGCTCAGACTACCTCTG 1707
Qy 306 ThrSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325
Db 1708 ACACTAATACTGCTACTCACCACCCAGCGCGCTGCTGCAAGCACCTTTCCGA 1767
Qy 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaIeuSer 345
Db 1768 CCTGCCCTTCACTTTAAACAAACCATACAGATGCTGCAATTGGAAGTGCACAGCCTTGAGC 1827
Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisIeu 365
Db 1828 GCCACTGAATCAGATGACTTTGGCCTTGTTTACTAGCCTAT---GTAAATTGTACACTTG 1884
Qy 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1885 TTTGGCCTGAGCTGGCAGTTGCAACCACTT---GGACAGATCTAT 1926
RESULT 7
LOCUS BQ735622 884 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT_8097117 NICHD XGC Emb4 Xenopus laevis cDNA clone
IMAGE:5542635 5', mRNA sequence.
ACCESSION BQ735622
VERSION BQ735622.1 GI:21874519
KEYWORDS EST.
SOURCE xenopus laevis (African clawed frog)
ORGANISM xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 884)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.lnl.gov/bbrp/image/image.html
Plate: LLAM12242 row: e column: 04
High quality sequence stop: 373.
FEATURES
Location/Qualifiers
1..884
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:5542635"
/dev_stage="embryo, stage 31-32"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD XGC Emb4"
/note="Organ: whole embryo; Vector: pCMV-SPORT6; Site: 1;
NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 2.1 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection
(XGC) library."

Qy 267 GluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheIeu 286
Db 1639 -----CATTCCTG 1647
Qy 287 PheIysProGly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 1648 TTCAACATGATGCTGTTCTTCTGCTATCTTCTGTCAGCCAGCTCAGACTACCTCTG 1707
Qy 306 ThrSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325
Db 1708 ACACTAATACTGCTACTCACCACCCAGCGCGCTGCTGCAAGCACCTTTCCGA 1767
Qy 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaIeuSer 345
Db 1768 CCTGCCCTTCACTTTAAACAAACCATACAGATGCTGCAATTGGAAGTGCACAGCCTTGAGC 1827
Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisIeu 365
Db 1828 GCCACTGAATCAGATGACTTTGGCCTTGTTTACTAGCCTAT---GTAAATTGTACACTTG 1884
Qy 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1885 TTTGGCCTGAGCTGGCAGTTGCAACCACTT---GGACAGATCTAT 1926
RESULT 7
LOCUS BQ735622 884 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT_8097117 NICHD XGC Emb4 Xenopus laevis cDNA clone
IMAGE:5542635 5', mRNA sequence.
ACCESSION BQ735622
VERSION BQ735622.1 GI:21874519
KEYWORDS EST.
SOURCE xenopus laevis (African clawed frog)
ORGANISM xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 884)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.lnl.gov/bbrp/image/image.html
Plate: LLAM12242 row: e column: 04
High quality sequence stop: 373.
FEATURES
Location/Qualifiers
1..884
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:5542635"
/dev_stage="embryo, stage 31-32"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD XGC Emb4"
/note="Organ: whole embryo; Vector: pCMV-SPORT6; Site: 1;
NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 2.1 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection
(XGC) library."

Qy 267 GluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheIeu 286
Db 1639 -----CATTCCTG 1647
Qy 287 PheIysProGly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 1648 TTCAACATGATGCTGTTCTTCTGCTATCTTCTGTCAGCCAGCTCAGACTACCTCTG 1707
Qy 306 ThrSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325
Db 1708 ACACTAATACTGCTACTCACCACCCAGCGCGCTGCTGCAAGCACCTTTCCGA 1767
Qy 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaIeuSer 345
Db 1768 CCTGCCCTTCACTTTAAACAAACCATACAGATGCTGCAATTGGAAGTGCACAGCCTTGAGC 1827
Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisIeu 365
Db 1828 GCCACTGAATCAGATGACTTTGGCCTTGTTTACTAGCCTAT---GTAAATTGTACACTTG 1884
Qy 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1885 TTTGGCCTGAGCTGGCAGTTGCAACCACTT---GGACAGATCTAT 1926
RESULT 7
LOCUS BQ735622 884 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT_8097117 NICHD XGC Emb4 Xenopus laevis cDNA clone
IMAGE:5542635 5', mRNA sequence.
ACCESSION BQ735622
VERSION BQ735622.1 GI:21874519
KEYWORDS EST.
SOURCE xenopus laevis (African clawed frog)
ORGANISM xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 884)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.lnl.gov/bbrp/image/image.html
Plate: LLAM12242 row: e column: 04
High quality sequence stop: 373.
FEATURES
Location/Qualifiers
1..884
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:5542635"
/dev_stage="embryo, stage 31-32"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD XGC Emb4"
/note="Organ: whole embryo; Vector: pCMV-SPORT6; Site: 1;
NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 2.1 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection
(XGC) library."


```

101 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
496 CGGAGCTGGTCTTGAAGATGTGACTCCACCCTGGGACTTTGTACCGGACTGATATT 555
101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
556 GGGCTGCTCACTGTGGCTACTCCATCAGCGCTGGCTCAGATGCCGACACGGAGCGGAT 615
121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
616 GTGTGATGTGCTCAGCTGAGCATCTCTGTGAGGCTCTGGGAGCGACACCAATCCGACGC 675
141 SerSer-CysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisG1 160
676 AGCTCCCTGCTGTGAGCGCGGC-AACTCCAACCTCACTC-ACCGACACGGAGCACGA 733
160 uAsnThrGluThr 164
734 GAATACCGAAACT 746

RESULT 9
BU119163
LOCUS
DEFINITION
603143046F1 CSEQCHL16 Gallus gallus cDNA clone ChEST137111 5', mRNA
sequence.
ACCESSION
BU119163
VERSION
BU119163.1 GI:25328007
KEYWORDS
EST.
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 887)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES
source
1..887
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="ChEST137111"
/sex="Female"
/tissue_type="not cerebrum or cerebellum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHL16"
/note="Organ: brain; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)
[Stratagene] vector to accommodate cDNA produced with the
T-trimmed protocol (Construction of uni-directionally
cloned cDNA libraries from messenger RNA for improved 3'
end DNA sequencing by Glenn Fu, et al. U.S. Patent #
6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI.
Ligate in double stranded adaptor containing BspI and
BamHI sites [5'ggcgcgtgcagcccgatccgaaaaag]
[5'aattcttttttcggtacggggtgcgcg]"

ORIGIN

```

```

Alignment Scores: 4.68e-40 Length: 887
Pred. No.: 687.50 Matches: 146
Score: 87.43% Conservative: 14
Percent Similarity: 79.78% Mismatches: 21
Best Local Similarity: 32.08% Indels: 5
Query Match: 13 Gaps: 0
DB:

US-10-029-020-14_COPY_1_400 (1-400) x BU119163 (1-887)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 343 ATGATGTAAAGAAAGAAACCGTATCGATCTCTGACTCGCGCGCGACACGAGCGC 402
QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 403 CGCTACACCACTCTTTCAGCGGAGAGTGAGGACAGAGGCTCTCTCAGAGTCTTATAGC 462
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 463 TCCAGTGAGACCTTGAAGGCTTATGATCAAGACTCCAGGTTGACCTACAGCAATCGGTC 522
QY 61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 523 AAAGACATGGTGCCACGAGAGGCTGATGATTCCTCGGAGCAGGAGCCAACTTCTCTTTG 582
QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 583 CGGAGCTGGGCTTTGAAGATGTGACTCCACCCATGGGACTTTGTACCGGACTGATATT 642
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 643 GGGCTGCTCACTGTGGCTACTCCATCAGCGTGGCTCAGATGCCGACACGGAGCGGAT 702
QY 121 ThrValLeuSer-ProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyAr 140
Db 703 GTGGTCAATGTACCCCTGAGCATCTCTGAGGCTCTCGGAGCGCACACCAAAATCCGAGC 762
QY 140 gSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisG1 160
Db 763 CAGCTCTGCTGTGTCGAGCGCGGCGCAACTCCAACTC-ACCCCTCACCGACACGGAAGCCCG 821
QY 160 uAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArg-ThrProp 180
Db 822 AGAATACGAAT-GATCATCTCTCCATCTTCA-AATCATCAAGACTCCGAATCCACAC 879
QY 180 roPro 181
Db 880 CTCCA 884

RESULT 10
CB544750
LOCUS
DEFINITION
CB544750 trypl-00009-a5 5', mRNA sequence.
ACCESSION
CB544750
VERSION
CB544750.1 GI:29428647
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 402)
AUTHORS
Amgen EST Program.
TITLE
Amgen Rat EST Program
JOURNAL
Unpublished (2003)
COMMENT
Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00009 row: a column: 5.
Location/Qualifiers

FEATURES

```



```

source      1. .402
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            /clone="tryp1-00009-a5"
            /tissue_type="prostate"
            /clone_lib="tryp1 (10582)"
            /note="Vector: pYVA-41l; Site_1: HindIII; Site_2: NotI;
            rat prostate"

ORIGIN
Alignment Scores:
Pred. No.:      2,29e-40      Length:      402
Score:          685.00      Matches:      126
Percent Similarity: 93.22%      Conservative: 2
Best local Similarity: 97.67%      Mismatches: 1
Query Match:      31.96%      Indels:      0
DB:              14          Gaps:          0

US-10-029-020-14_COPY_1_400 (1-400) x CB544750 (1-402)
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 2 AGAAACCTAGCAGCAGCAGCCTTCTAGGACATTTCAGGACATTCAGGACCAACCTCATTGAGATGGAC 61
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 62 ATTCTCAGCGCTCCCGCATGATGGGCGCTTACAGTGACGGGCACCTTCTCTTCAAGCCC 121
QY 290 GlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThr 309
Db 122 GGAGCGACCTCCCGCAGCTTCTGACACACAGTCCCGAGGATACCCCTAAACGCTAGCACC 181
QY 310 ValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsn 329
Db 182 GTGTATTGCGCCCGCCCGCCCTGCGCGAGCAGCCTTTTCGAGACGACGCTTTAAAC 241
QY 330 LeuLysLysProSerLysTyrCysAsnTrpLysCysAlaLeuSerAlaIleValIle 349
Db 242 CTCAGAAGCGCTTCCAAGTACTGCACTGGAAGTGGCGGCACTAAGCGGCATCTCATC 301
QY 350 SerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMethHisLeuPheGlyLeuAsn 369
Db 302 TCAGCTACGCTCGTCATCTGCTGCGATACCTTGTGGCCATGCACCTGTTGGCCTAAAC 361
QY 370 TrpHisLeuGlnProMetGluGlyGln 378
Db 362 TGGCACCTGACGCGATGGAGGGGCGAG 388

RESULT 11
AK034286
LOCUS      4556 bp      mRNA      linear      HTC 18-SEP-2003
DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length
            enriched library, clone:9330172M18 product:odd Oz/ten-m homolog 1
            (Drosophila), full insert sequence.
ACCESSION  AK034286
VERSION    AK034286.1  GI:26083876
KEYWORDS   HTC; CAP trapper.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1
AUTHORS    Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE      Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
            10349636
MEDLINE    99279253
PUBMED     10349636
REFERENCE  2
AUTHORS    Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE      Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
            10349636
JOURNAL    MEDLINE
PUBMED     11042159
REFERENCE  3
AUTHORS    Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
            Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
            Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
            Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
            Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
            Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
            RIKEN integrated sequence analysis (RISA) system-384-format
            sequencing pipeline with 384 multipillar sequencer
            Genome Res. 10 (11), 1757-1771 (2000)
            20530913
            11076861
JOURNAL    MEDLINE
PUBMED     11076861
REFERENCE  4
AUTHORS    The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
            Functional annotation of a full-length mouse cDNA collection
            Nature 409, 685-690 (2001)
            5
JOURNAL    MEDLINE
PUBMED     11076861
REFERENCE  5
AUTHORS    The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
            Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
            6 (bases 1 to 4556)
JOURNAL    MEDLINE
PUBMED     11076861
REFERENCE  6
AUTHORS    Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
            Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
            Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
            Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T.,
            Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
            Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
            Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
            Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
            Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
            Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S.,
            Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
            Muramatsu, M. and Hayashizaki, Y.
            Direct Submission
            Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
            CNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/
FEATURES
            Location/Qualifiers
            1. 4556
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="FANTOM DB:9330172M18"
            /db_xref="MGI:2398485"
            /db_xref="taxon:10090"
            /clone="9330172M18"
            /sex="male"
            /tissue_type="dienecephalon"
            /clone_lib="RIKEN full-length enriched mouse cDNA library"
            /dev stage="adult"
            1. 4556
            /note="odd Oz/ten-m homolog 1 (Drosophila)
            (MGI:MGI:1345185, GB|NM_011855, evidence: BLASTN, 99%,
            match=3037)"
            misc_feature
            1. 4556
            /note="odd Oz/ten-m homolog 1 (Drosophila)
            (MGI:MGI:1345185, GB|NM_011855, evidence: BLASTN, 99%,
            match=3037)"

```



```

US-10-029-020-14_COPY_1_400 (1-400) x AL922332 (1-490)
QY 3 ValysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArgTyr 22
Db 489 GTCAAGGACGACGACCTACCGCTCTCTGACCTCCAGGCGGACACGAGCGCGGTAC 430
QY 23 ThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSerSer 42
Db 429 ACCAGCTCATCTCCGACACGAGGATGCGCAAGATCAACCTAAATCTTACGCTCGAGT 370
QY 43 GluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgValLysAsp 62
Db 369 GAGACCTCAAGCCTTCGACACGAGACTCCAGACTGGCTATGCGACGCGGTCAAGAC 310
QY 63 IleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeuArgGlu 82
Db 309 CTGGTGACCATGAGCGCGACGAGTTCAGCAGACGAGCGCGACTTTTCTCTCAGAGAC 250
QY 83 LeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIleGlyLeu 102
Db 249 ATGGCTTGGGATCCCGCGCCACATATGGCAGCGTACCGACAGAAATGGCCCTT 190
QY 103 ProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrVal 122
Db 189 CCCACCGGACTACTCAGTGTAGTGTGGCATCAGACGCTGACACAGAAACACAGCGCAT 130
QY 123 LeuSerProGluHisProValArgLeuThrGlyArgSer---ThrArgSerGlyArgSer 141
Db 129 ATGTCCCGGACGCGCGTCCCGCTCTGGGCGCGCAGCAACACCAATCCGCGCGCAGT 70
QY 142 SerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsn 161
Db 69 TCCTGCTGTCCAGCAGAGCAACTCCAACTTACGCTCAGCAGCAGCAGCATGAAAC 10
QY 162 ThrGlu 163
Db 9 ACTGAG 4

BG036207 870 bp mRNA linear EST 24-JAN-2001
602326960F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4428351 5',
RNA sequence.
BG036207 1 GI:12431132
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTP/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10178 row: h column: 16
High quality sequence stop: 713.
Location/Qualifiers
1. .870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4428351"
/tissue_type="adenocarcinoma, cell line"

FEATURES
source
RESULT 13
BG036207
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AY405291 511 bp DNA linear GSS 12-DEC-2003
Homo sapiens HCM2175 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY405291
AY405291.1 GI:39761265
GSS.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 511)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 511)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering

```

```

/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"

```

```

/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2,39e-31 Length: 870
Score: 568.00 Matches: 109
Percent Similarity: 99.09% Conservative: 0
Best Local Similarity: 99.09% Mismatches: 1
Query Match: 26.50% Indels: 1
DB: 10 Gaps: 0

```

```

US-10-029-020-14_COPY_1_400 (1-400) x BG036207 (1-870)

```

```

QY 291 GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrVal 310
Db 2 GGCACCTCCCGCTTTCTGACCAACATCACCAGG-TACCCACTGACGTCCAGCACAGTG 60
QY 311 TyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeu 330
Db 61 TACTCTCTCCGCGCGCCGACCTTCCGCGCGAGCCTTCCGCGCGCGCTTTTAACTTC 120
QY 331 LysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSer 350
Db 121 AAGAGCCCTCAAGTACTGTAAGTGGAGGCGCAGCCCTGAGCGCATCGTATCTCA 180
QY 351 AlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrp 370
Db 181 GCCACTCTGGTTCATCTCTGCGCATCTTTGTGCGCATGACCTGTTGGCTAACTGG 240
QY 371 HisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSerSerTrp 390
Db 241 CACCTGAGCGCATGGAGGGGCGAGATGATGAGATCAGGAGGACACAGCGAGGTGG 300
QY 391 ProValProThrAspValSerLeuTyrPro 400
Db 301 CCTGTCCACACGACGCTCTCCCTATATACCC 330

```

RESULT 14

```

AY405291
LOCUS

```

```

DEFINITION

```

```

ACCESSION

```

```

VERSION

```

```

KEYWORDS

```

```

SOURCE

```

```

ORGANISM

```

```

REFERENCE

```

```

AUTHORS

```

```

TITLE

```

```

JOURNAL

```

```

PUBMED

```

```

REFERENCE

```

```

AUTHORS

```

```

TITLE

```

```

JOURNAL

```

```

COMMENT

```

Inferring nonneutral evolution from human-chimp-mouse orthologous

TITLE
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
2 (bases 1 to 511)
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..511
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>511
/locus_tag="HCM2175"
gene
ORIGIN
Alignment Scores: 5,13e-29 Length: 511
Pred. No.: 531.50 Matches: 108
Score: 79.88% Conservative: 23
Percent Similarity: 79.88% Mismatches: 30
Best Local Similarity: 65.85% Indels: 3
Query Match: 24.80% Gaps: 3
DB:
US-10-029-020-14_COPY_1_400 (1-400) x AY405292 (1-511)
QY 1 MetAspVallysGluArglyProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 1 ATGGATGTGAAGAACGAGCGCTTACTGCTCTCTGACCAAGAGCAGACGAGAGGAA 60
QY 20 ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 61 CGCGCTACACAAATTCCTCCGAGCAATGAGAGTGCGGGTACCCACACAGAGTCC 120
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 121 TACAGTTCCAGCAGACATGAAAGCTTTTGTATCATGATTCCTCGCGCTGCTTACGGC 180
QY 58 SerArgVallysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsn 77
Db 181 AACAGAGTGAAGGATTTGGTTTCACAGAGAAGCAGAGTTCACAGAGACAGAAAT 240
QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
Db 241 TTTACCTTAAGCAGTAGGAGTTTGTGACCACTCGAACCACTCGAAGAGGACTGGCATTTGT 300
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 301 CGGGAATGGGCTCCCTCAGAGAGTTTACTATCATGTCAGTGGGTCAGATGCTGATACT 360
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137
Db 361 GAAATGAAGCAGTGATGTCCTCCAGAGCATGCCAGACTTTGGGGCAGGGGGTCAAA 420
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 421 TCAGGCCGAGCTCTGCTGCTGTCAGTGGTCCAACTCAGGCCTCAGCCTGACAGATACG 480
QY 158 GluHisGluAsn 161
Db 481 GAGCAGGAAAC 492

Search completed: August 14, 2004, 18:02:37
Job time : 4710.12 secs

FEATURES
source
Location/Qualifiers
1..511
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>511
/locus_tag="HCM2175"
gene
ORIGIN
Alignment Scores: 5,13e-29 Length: 511
Pred. No.: 531.50 Matches: 108
Score: 79.88% Conservative: 23
Percent Similarity: 79.88% Mismatches: 30
Best Local Similarity: 65.85% Indels: 3
Query Match: 24.80% Gaps: 3
DB:
US-10-029-020-14_COPY_1_400 (1-400) x AY405291 (1-511)
QY 1 MetAspVallysGluArglyProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 1 ATGGATGTGAAGAACGAGCGCTTACTGCTCTCTGACCAAGAGCAGACGAGAGGAA 60
QY 20 ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 61 CGCGCTACACAAATTCCTCCGAGCAATGAGAGTGCGGGTACCCACACAGAGTCC 120
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 121 TACAGTTCCAGCAGACATGAAAGCTTTTGTATCATGATTCCTCGCGCTGCTTACGGC 180
QY 58 SerArgVallysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsn 77
Db 181 AACAGAGTGAAGGATTTGGTTTCACAGAGAAGCAGAGTTCACAGAGACAGAAAT 240
QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
Db 241 TTTACCTTAAGCAGTAGGAGTTTGTGAAACCACTCGAACCACTCGAAGAGGACTGGCATTTGT 300
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 301 CGGGAATGGGCTCCCTCAGAGAGTTTACTATCATGTCAGTGGGTCAGATGCTGATACT 360
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137
Db 361 GAAATGAAGCAGTGATGTCCTCCAGAGCATGCCAGACTTTGGGGCAGGGGGTCAAA 420
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 421 TCAGGCCGAGCTCTGCTGCTGTCAGTGGTCCAACTCAGGCCTCAGCCTGACAGATACG 480
QY 158 GluHisGluAsn 161
Db 481 GAGCAGGAAAC 492
RESULT 15
AY405292 511 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HCM2175 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY405292
VERSION AY405292.1 GI:39761266
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 511)
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.